Protocol for Ozone Modeling of the Houston/Galveston/Brazoria Area:

Combined 1- and 8-hour Ozone Modeling Analysis

Prepared by the

Technical Support Section
Technical Analysis Division
Office of Environmental Policy, Analysis & Assessment
Texas Natural Resource Conservation Commission

February, 2004

Table of Contents

List of Figures

List of Tables

1 Summary

This protocol presents procedures the Texas Commission on Environmental Quality (TCEQ) will use to model ozone in the Houston/Galveston/Brazoria (HGB) area using an approved photochemical model. The focus of this modeling is two-fold: First, we will conduct one-hour attainment demonstration analyses to complete the second phase of the Mid-Course Review (MCR), and second we will conduct early modeling for the eight-hour ozone standard using a future year of 2010. The modeling will include the episode modeled in Phase 1 of the MCR, plus additional episode days from the period of the 2000 Texas Air Quality Study (TexAQS 2000, or simply TexAQS). This expanded episode will help ensure that control strategies (adopted or proposed) will be effective over the most prevalent meteorological conditions associated with the formation of unhealthy levels of ozone in the region. As in the Phase 1 MCR modeling, this analysis will rely heavily upon data collected during the TexAQS 2000, and will incorporate recent scientific advancements as appropriate. Because this modeling combines the one-hour analysis planned for Phase 2 of the MCR with initial modeling for the eight-hour ozone standard, we refer to this round of modeling as the Combined 1- and 8- hour Ozone Modeling Analysis, or COMA.

The objective of this modeling protocol is to maintain and enhance the technical credibility of the study by establishing in advance agreed-upon procedures for conducting a successful modeling project. Section 2 of the protocol describes the background, objectives, schedule, and organizational structure of the study. The remainder of the protocol describes the structure of the modeling system, the development of needed model databases, the plans for model performance evaluation, the procedures to be used to determine whether proposed control strategies are sufficient to show attainment of the eight-hour ozone standard, and the procedures for documenting the study results. This protocol also includes conceptual models for both one and eight-hour ozone formation in the HGB airshed (as Appendices A and B). The conceptual model provides a qualitative description of the region's ozone problem and the many factors which collectively result in exceedances of the two ozone standards.

Current plans are to submit a SIP revision based on this modeling to EPA in Fall, 2004. The SIP revision is intended to satisfy any remaining requirements for a one-hour ozone demonstration and to develop a roadmap for achieving attainment of the eight-hour ozone standard by the area's attainment date. At this time, the actual eight-hour attainment date for the HGB area is uncertain, since EPA has not issued final Guidance for implementation of the eight-hour standard.

A second, but potentially even more important goal of this modeling is to aid in advancing the understanding of the many complex processes and interactions that cause ozone exceedances along the upper Texas coast. The TexAQS 2000 has yielded an immense set of air quality, meteorological, and emissions data which has significantly advanced the science of ozone air pollution in Texas and elsewhere. The modeling provides a means of integrating all the disparate elements of the TexAQS 2000 study into a holistic three-dimensional picture of the HGB airshed necessary to study the interplay of the many factors which drive Houston's ozone problem.

This protocol reflects the current plans of the TCEQ modeling staff but may be modified to account for new science, better modeling tools, changes in resources, or other events. This protocol should be considered a living document which changes as necessary to reflect the current plans of the TCEQ, in coordination with EPA Region VI.

2 Ozone Modeling Study Design

2.1 Background and Objectives

The 1990 Federal Clean Air Act (FCAA) Amendments established five classifications for ozone nonattainment areas based on the magnitude of the monitored one-hour ozone design values, and established dates by which each classified area should attain the standard. The Houston/Galveston/Brazoria Consolidated Metropolitan Statistical Area (Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, and Waller Counties) was designated as a severe-17 ozone nonattainment area, with an attainment year of 2007. EPA's proposal for implementation of the eight-hour standard translates these classifications to apply to eight-hour design values. Under the proposal, the HGB area would be classified as either Moderate or Serious, depending upon ozone levels recorded in 2003. Since EPA has not at this writing completed the Eight-Hour Implementation Guidance nor made its designations, it is impossible to know what the area's eight-hour attainment year will be, but it almost certainly will not be earlier than 2010. Since this year is expected to coincide with the attainment year for other Texas areas, we have selected this as the year to use in the early eight-hour assessment.

Under the one-hour NAAQS, the HGB area is classified as Severe-17 and must attain the one-hour ozone NAAQS by 2007.

There have previously been five distinct rounds of the HGB attainment demonstration modeling, dating back to 1990. Because the first four rounds have little relevance to current work, discussions of these rounds have been deleted from this version of the modeling protocol. However, the initial modeling for the eight-hour ozone standard is largely an extension of the previous round of modeling (Phase 1 of the MCR), which is described below.

<u>2.1.1 Phase 1 of the Mid-Course Review and the TexAQS 2000</u>

The modeling for Phase 1 of the Mid-Course Review represented a significant departure from previous work. This was the first modeling to incorporate the TexAQS 2000 data and - more significantly - the improved understanding of ozone formation in the region gained through analysis of this data. The TexAQS provided copious amounts of meteorological, air quality, and emissions data which were used to develop modeling that characterized the region's ozone formation processes much better than was possible previously.

In this round of modeling, the meteorological modeling was conducted using the state-of-the-science PSU-NCAR MM5 model. The photochemical modeling was conducted with CAMx Version 3.1. Almost the entire modeling inventory used in the Phase 1 modeling was built from the ground-up, and many emission categories changed significantly from previous modeling. A special emissions inventory was conducted which provided hourly emissions from many of the region's largest point sources.

Data collected during the TexAQS 2000 were incorporated into the meteorological modeling performed with MM5 to the extent possible in time to meet the schedule for Phase 1 of the MCR. Specifically, five radar profilers were operating during the study, and data from these were used to nudge MM5 winds. Additional data, especially aircraft observations, were used to help characterize vertical mixing. Possibly the most significant finding of the TexAQS is that reported emissions of light olefins (and possibly other hydrocarbons) are not consistent with measured atmospheric concentrations. As a result of these findings, TCEQ adjusted emissions of several highly-reactive VOCs in the modeling. As noted earlier, scientific analyses of the TexAQS 2000 study results were incorporated into the modeling process whenever appropriate.

The Phase 1 MCR modeling was conducted specifically to address the issue "Can VOC emission reductions be substituted for part of the point source NO_X reductions adopted in 2000?". Future year 2007 modeling was performed to address this issue, and the TCEQ concluded that it was indeed possible to substitute reductions of emissions of highly-reactive VOCs for the last 10% of NO_X reductions. On December 13, 2002, the TCEQ adopted a SIP revision which adjusted the NO_X emission reductions required for industrial sources from (nominally) 90% to 80%. The SIP revision also included significant reductions in emissions of a number of highly-reactive VOCs (ethylene, propylene, 1,3 butadiene, and all butenes). Modeling analyses showed that the combined VOC/NO_X strategy would produce much greater air-quality benefits than the previous NO_X-only strategy, but that additional emission reductions would still be required to reach attainment.

2.1.2 Combined 1- and 8-hour Ozone Modeling Analysis (COMA)

The COMA will both enhance and expand upon the modeling performed for Phase 1 of the MCR. The Phase 1 episode has been extended to include nine additional days. In this round, TCEQ plans to incorporate a number of additional enhancements to the modeling inventory. Chief among these enhancements will be refined adjustments to certain highly-reactive VOC emissions based on the results of several studies which are now underway or planned. Modeling will also be conducted to assess the effects of adjustments to other, generally less-reactive, VOC emissions as well.

As in Phase 1, the meteorological modeling will be conducted using the state-of-the-science PSU-NCAR MM5 model. Several areas of potential improvement to the meteorological characterization were identified during Phase 1, and we are implementing several of these improvements. The most significant advance is the incorporation of soil heat capacity estimates derived from satellite measurements. Soil heat capacity is a critically important parameter in establishing wind fields and mixing depths, but heretofore has been largely guessed at, since almost no measurement data has been available. Wherever possible, additional TexAQS 2000 data will be incorporated into the modeling process. The photochemical modeling will be conducted with CAMx Version 4 (or a later version, if one becomes available).

2.2 The Texas Air Quality Study 2000

From August 15 to September 15, 2000, approximately 250 investigators from more than 35 organizations joined the TCEQ in the TexAQS 2000 to carry out field research to improve technical understanding of the factors affecting ozone and fine particle concentrations in the eastern half of Texas. TexAQS 2000 was based in Houston, and its work concentrated on the Houston region. TexAQS 2000 collected extensive data useful for supporting photochemical modeling of episodes that occurred during the study period. Figure 1 shows the maximum one-hour ozone concentration each day of the field study, and it also shows the number of continuous monitoring stations that reached or exceeded 125 parts per billion (ppb) each day. Figure 2 shows the maximum eight-hour ozone concentrations and the number of continuous monitoring stations that reached or exceeded 85 ppb.

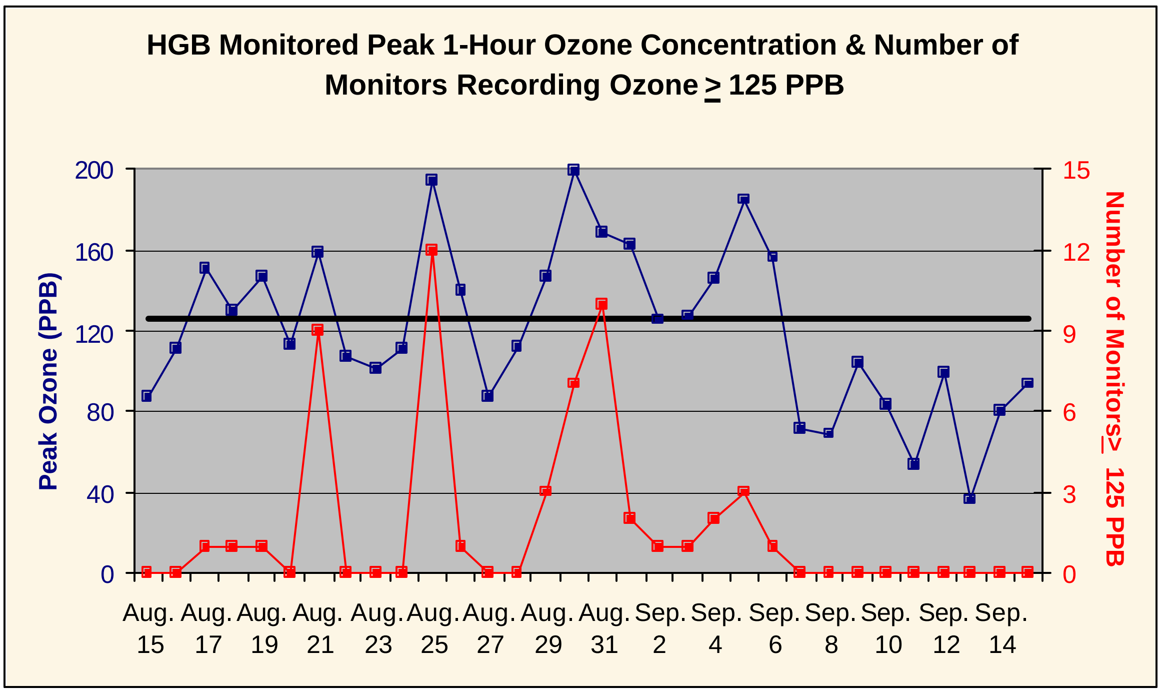

Figure 1: One-hour peak ozone concentrations and number of monitors recording one-hour ozone concentrations ≥ 125 parts/billion during TexAQS.

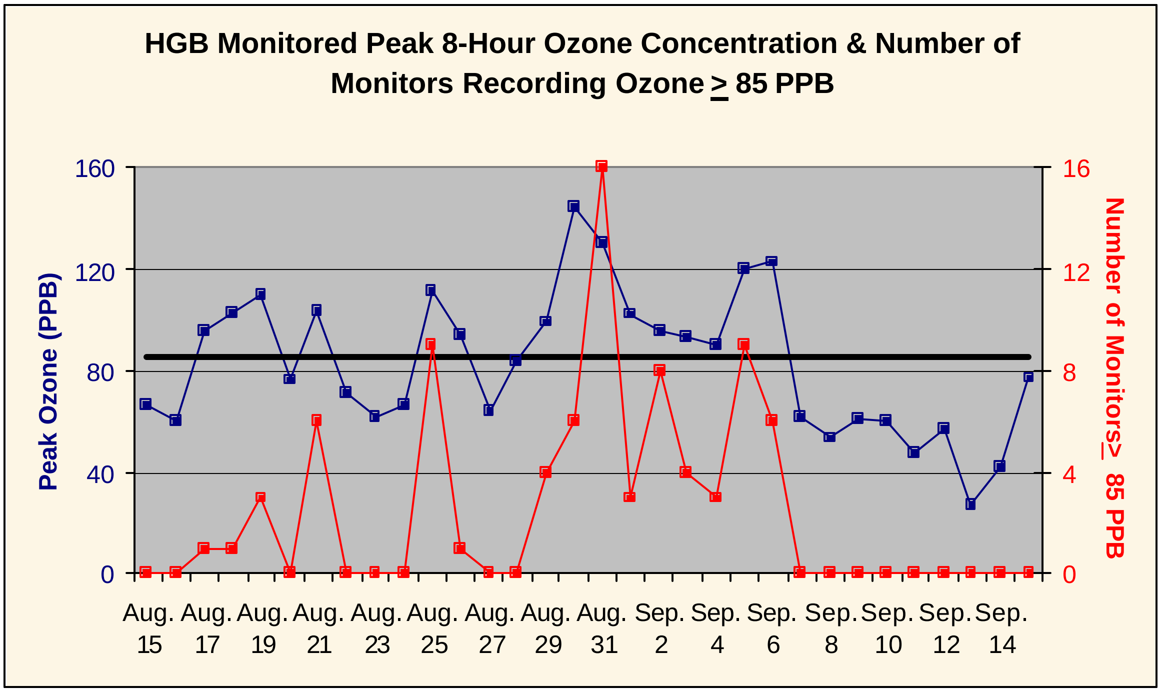

Figure 2: Eight-hour peak ozone concentrations and number of monitors recording eight-hour ozone concentrations ≥ 85 parts/billion during TexAQS.

The portion of the study period selected for modeling (the "modeling episode" - see Section 4.2) is August 18-September 6. Although some high 8-hour ozone concentrations were seen early in the episode, these exceedances were not widespread and are not of primary interest in the current modeling application (these days may be studied in more detail at a later time). Thus, August 18-20 is considered the ramp-up period for the main episode period. The remaining portion of the episode contains several shorter sub-episodes exhibiting several different types of behaviors.

- The first sub-episode was a widespread one-day exceedance (August 21) which featured high one- and eight- hour concentrations (159 and 104 parts/billion, respectively) with one-hour concentrations at eight monitors exceeding 125 parts/billion and with eight-hour concentrations above 85 parts/billion at six monitors.

- Following this day was a period of relatively low ozone which lasted until August 25. On August 25, twelve monitors recorded one-hour averages exceeding 125 parts/billion, and nine monitors recorded eight-hour averages above 85 parts/billion. August 25 was characterized by a fairly compact parcel of ozone-laden air which moved westward and brought very high ozone levels to some areas; the one-hour peak was 194 parts/billion. The eight-hour peak of 111 parts/billion was sizable, but not extraordinarily so.

- August 26 saw relatively modest exceedances of both the one- and eight-hour standard (140 & 94 parts/billion respectively), both of which occurred in Conroe to the north of the urban center.[1]

- August 27 and 28 saw relatively low ozone concentrations (although the eight-hour peak on August 28 was just below the standard at 84 parts/billion). Then beginning on August 29, both the one- and eight-hour standards were exceeded for nine consecutive days.

- On August 29, four monitors recorded eight-hour exceedances while three saw exceedances of the one-hour standard. Peak eight-hour ozone was 99 parts/billion, and the one-hour peak was 146 parts/billion.

- The next day, August 30, saw the highest one-hour concentration of the entire year, 199 parts/billion at the HRM-8 monitor in LaPorte. The eight-hour peak was very high, 144 parts/billion. Despite its intensity, the area of high ozone was not especially large, with six monitors exceeding the eight-hour standard and seven exceeding the one-hour standard.

- The following day, August 31 saw much more widespread ozone with sixteen monitors recording eight-hour peaks over 85 parts/billion and 10 recording one-hour concentrations above 125 parts/billion. Despite the broad area affected, peak one- and eight-hour concentrations, 168 and 130 parts/billion, respectively, were significantly lower than on the previous day.

- September 1 saw one-hour exceedances at only two monitors, and eight-hour exceedances at three as a west wind transported much of the city's pollution eastward. The one-hour peak was still fairly large at 163 parts/billion, but the eight-hour peak was a relatively moderate 102 parts/billion.

- September 2 and 3 had minimal one-hour exceedances, 125 and 127 parts/billion, respectively, at a single station on each day. Eight-hour exceedances, however, were more widespread with eight and four stations, respectively, exceeding the 85 part/billion standard. Peak eight-hour ozone on September 2 was 95 parts/billion, followed by 93 parts/billion on September 3.

- On September 4, three monitors exceeded the eight-hour standard with a peak of 90 parts/billion. But one-hour ozone was a little higher, peaking at 145 parts/billion at one of the

[1]Note that it is not technically correct to refer to individual measured eight-hour concentrations above 85 parts/billion as "exceedances", since the form of eight-hour standard is now a three-year average of fourth-highest concentrations. However, for convenience in this document, we will occasionally refer to 8-hour average concentrations in excess of 85 parts/billion as "exceedances".

two monitors recording exceedances on that day.

- On September 5, a sizable one-hour concentration of 185 parts/billion was recorded, but only three monitors reported one-hour exceedances. However, nine monitors reported eight-hour exceedances with a peak of 120 parts/billion.

- September 6 likewise recorded a relatively high one-hour peak of 156 parts/billion at the only station exceeding the one-hour standard. But again the eight-hour exceedance area was fairly widespread with six monitors reporting exceedances and a peak of 123 parts/billion.

Note that the modeling period includes several days when no regulatory monitor in the Houston area recorded a 125 ppb concentration. These days will provide a test of the model's performance when similar emissions did not produce high ozone concentrations at any of the area's monitors. A second consideration in extending the episode, especially in August, was to allow use of Automatic Gas Chromatograph measurements which was available during that time period. The Deer Park Auto-GC was lost due to mechanical failure on August 25, but prior to that time collected a significant amount of valuable hydrocarbon data which can be used in model evaluation.

2.2.1 Data Collection

Listed below are the major components of the TexAQS 2000.

- Six research aircraft, four of which were based in Houston and performed multiple missions as described:
  - The National Oceanic and Atmospheric Administration (NOAA) used a Lockheed Electra as a platform to collect regional chemistry and meteorological measurements to help define regional emissions, chemistry, and transport.
  - The Department of Energy provided a Gulfstream 1 with instrumentation similar to the Electra's to measure both regional and local emission, chemistry, and transport.
  - Baylor University operated a Twin Otter for the TCEQ, carrying advanced air quality monitoring instruments similar to those at a Level 2 ground station along with canisters for sampling volatile organic compounds. The Twin Otter's ability to fly slowly made it well-suited to studying urban and industrial plumes.
  - NOAA's Environmental Technology Laboratory provided a DC-3 aircraft to measure ozone and fine particles with a downward-looking LIDAR system well-suited to measuring the formation and movement of pollution plumes and to studying the effects of coastal meteorology, including the bay breeze.
  - NASA provided two aircraft for use in thermal mapping to help define and evaluate urban and industrial heat-island effects.

- Additional meteorological monitoring to provide data to help describe and understand how wind flows are influenced by bay breezes, sea breezes, and urban and industrial heat islands:

    - Six radar profilers to measure winds and virtual temperature aloft.
    - Two advanced acoustic sounders for the same purpose.
    - Three weather balloon launch sites to measure the temperature and moisture structure of the atmosphere, one of which also had global positioning system capability to measure winds aloft as well as the structure of the atmosphere.

- A Doppler LIDAR to aid in analysis of the interaction of the bay breeze and the industrialized Ship Channel area.

- To the approximately 50 routine, ground-based continuous ozone monitoring sites across the eastern half of Texas and neighboring states, the study added the following:

    - three Level 2 chemistry monitoring stations to provide detailed, high-sensitivity atmospheric chemistry information on ozone, sulfur dioxide, carbon monoxide, NO, and NO_2;
    - a principal atmospheric chemistry and physics research site at La Porte Airport at which many researchers from universities and national laboratories operated state-of-the-science instruments to investigate atmospheric processes and measure pollutant concentrations; and
    - a smaller advanced research site high on the Williams Tower, about 850 feet above ground level.

- An hour-by-hour emission inventory of emissions from the Houston industrial area reporting a much more detailed record of emissions than is normally available for either data analysis or photochemical modeling.

2.2.2 The Science Coordinating Committee

In August of 2001, the TCEQ hosted a four-day workshop to allow the TexAQS 2000 researchers to present their preliminary findings to each other, regulators, the regulated community, and the public. This extremely successful workshop provided air quality professionals with a wide variety of information on meteorology, atmospheric chemistry, and emissions, but posed many more questions than could be answered. During the workshop, a new group called the Interim Science Coordinating Committee (ISCC) was formed to provide a forum for researchers and other air quality professionals to continue to share information and to help guide additional research efforts required to assimilate the vast quantities of data into a broad understanding of the causes and potential remedies for the HGB region’s air quality problems. The ISCC, since renamed to the Science Coordinating Committee (SCC) is headed by an Executive Committee drawn from a larger Steering Committee. Several workgroups operate under the SCC to study specific air quality issues identified during the August, 2000 workshop. These workgroups are: Air Quality Meteorology and Modeling, Chemistry, Data Exchange, Emissions Characterization, Monitoring/Ambient Measurement and Mobile Emissions. In addition to the workgroups, a Synthesis Team was formed to assimilate information from the various groups into a coherent picture and to produce one or more Findings Documents summarizing the results.

The TCEQ maintains web pages devoted to Houston/Galveston Air Quality Science Evaluation, which contains information on the SCC and its various workgroups. This information can be accessed at http://www.tnrcc.state.tx.us/air/aqp/airquality_science.html

2.2.3 SCC Science Projects

The primary task of the Science Coordinating Committee (SCC) and its various workgroups was to help develop a list of projects that will improve the understanding of the Houston/Galveston ozone pollution problem, and to evaluate these projects in terms of their technical merit. Funding of projects by TCEQ and other organizations is guided, at least in part, by the recommendations of this committee.

Fiscal Year 2004 Projects

There are many Fiscal Year 2004 candidate projects designed to contribute to refinements in meteorological and photochemical modeling. A list of these projects is maintained at the following web site, which is updated periodically:
http://www.tnrcc.state.tx.us/air/aqp/airquality_impscience.html

Three of the most important questions for air quality research work this fiscal year (FY-2004) are:

- How can we best refine the emission adjustments used for olefins and for other classes of reactive compounds?

- What additional classes of compounds play significant roles in ozone formation in the HGB area, and are controls needed for these additional classes?

- What are the boundary conditions during the 2000 modeling episodes for the Houston and Beaumont-Port Arthur areas and the 1999 modeling episode for the DFW area and the near-nonattainment areas?

The first two of these questions are closely related and are addressed by several proposed projects including:

Project 46, “Source attribution and emission adjustment project”

Project 43, “More extensive analysis of four TexAQS 2000 high ozone days” and

Project 40, “Analysis of unusual (besides light olefins) VOCs observed in Houston”

One task specifically addresses the last question, although the TCEQ modeling staff plan to address this issue separately in the coming months:

Project 101, "Southern Oxidants Study (SOS) state-of-the-science assessment of information on transport and modeling episode boundary conditions". This project will improve estimates of boundary conditions for the modeling episodes and will aid in planning for the study of transport and of boundary conditions during the 2005-2006 field study.

Many other important projects under consideration are also listed, although it is probable that many can not be funded this year. Interested parties should contact Dr. Jim Price at jprice@tceq.state.tx.us or Laurel Carlisle at lcarlisl@tceq.state.tx.us.

2.3 Schedule

Table 1 shows projected milestones for the COMA. Detailed discussions of most of the items listed can be found later in this document.

Note that the previous version of this protocol had included plans to model two additional episodes (August 1-5 and 26-30, 1998), but now these episodes have been shelved in favor of extending the August-September 2000 TexAQS episode. The decision to focus on the TexAQS period was made because the extremely rich set of ambient data available during this period allows for very comprehensive model validation and also allows the model to be used as an aid in characterizing the exceedingly complex physico-chemical processes leading to ozone exceedances in the area.

Table 1 - COMA Milestones and Schedule

Milestones	**Completion Date**
Develop modeling protocol	Completed
Modeling of the August 18 - September 6, 2000 Episode:	
Base case emissions development:	
Develop TexAQS Special Inventory for additional days	Completed
Assess drought & heat stress model for biogenic emissions	Completed
Enhanced area/nonroad emissions	Completed
Develop 1-km inventory for fine-grid subdomain	Completed
Extend modeling inventory to cover extended episode	Completed
Point source hydrocarbon emission adjustment factor development	Ongoing

Milestones	**Completion Date**
Model-ready base case inventory (Carbon-Bond IV) completed	Completed
Update base inventory with new on-road emissions from HGAC/TTI	Completed
Meteorological model development:	
Model extended episode with MM5	Completed
Incorporate satellite-derived soil heat capacity	Completed
MM5 modeling completed	Completed
Base case model performance evaluation completed	Completed
2007 future inventory developments completed	Completed
2007 future case modeling with Phase 1 controls completed	Completed
2007 future case attainment modeling completed	January, 2004
2010 future inventory developments completed	January, 2004
2010 eight-hour assessment modeling completed	February, 2004
Documentation completed	March, 2003
Proposed SIP Revision	April, 2003
SIP Adoption	October, 2004

2.4 Modeling Policy Oversight Groups

The Regional Air Quality Planning Committee provides oversight and review of photochemical modeling as related to development of the SIP and policy implications. This group is coordinated by the Houston-Galveston Area Council, the Metropolitan Planning Organization for the HGB area. The membership of this committee can be found at http://www.h-gac.com/HGAC/Departments/Transportation/Committees/RAQPC/Regional_Air_Quality_Planning_Committee+.html.

2.5 Photochemical Modeling Technical Committee

The Photochemical Modeling Technical Committee serves in an advisory role for the technical aspects of applying photochemical modeling and improving the science. The members of this committee are listed in Table 2.

Table 2. Photochemical Modeling Technical Committee

NAME	REPRESENTING
Dave Allen	University of Texas at Austin
Ramon Alvarez	Environmental Defense Fund
Dan Baker	Equilon Enterprises LLC
Rob Barrett	Harris County Pollution Control Division
Harless Benthul	Benthul, Kean & Woodruff
Pamela Berger	Mayor’s Office, City of Houston
Craig Beskid	National Urban Air Toxics Research Center
Daewon Byun	Department of Geosciences, University of Houston
Jose Campos	FHWA
Chris Colville	Trinity Consultants
Hsing-wei Chu	Lamar University
Walter Crow	Radian
Alex Cuclis	Environmental Institute of Houston
Mike Cybulski	Clean Air Engineering
Weiping Dai	Trinity Consultants
Stephen Davis	TCEQ
Doug Deason	ExxonMobil Chemical Company
John Dege	DuPont
Tom Diggs	EPA
Jon Fisher	Texas Chemical Council
Richard Flannery	TCEQ, Region 12
Candy Garrett	TCEQ
Monica R. Gaudet	Metropolitan Transit Authority
Joseph Goldman	CLEAN & ICSEP
Reza Golkarfard	HGAC
Dennis Griffith	RAQPC
K Hackett	HGAC
John Hall	TERC
Alan Hansen	EPRI
Al Hendler	URS Corporation
Liz Hendler	Mid-Course Coalition
T. F. Henken	Baytown
April Hinson	DuPont
David Hitchcock	HARC
Thomas Ho	Lamar University
Ileana Isern-Flecha	TCEQ
Robert E. James	TCEQ, Region 12
S.C. Kilpatrick	Dow
Alan J. Krol	Amoco
John Kush	Reliant Energy
Jane Laping	City of Houston
Carole Lenz	Commissioner Radack, Harris County, RAQPC, H-GAC
Jacqueline Lentz	City of Houston
Jim Lester	Houston Advanced Research Center
Graciela Lubertino	EPA
Fred Manhart	SETRPC

Gene McMullen	Bureau of Air Quality Control, City of Houston
Susan Moore	BP
Quang Nguyen	EPA Region VI
Eduardo Olaguer	HARC
Barbara Pederson	DuPont
Charles E. Pehl	Pehl Environmental Consulting
Karl Pepple	HGAC
Chris Rabideau	Shell
Rebecca Rentz	Bracewell & Patterson
Dick Robertson	Texas Utilities
David Schanbacher	TCEQ
Charles Schleyer	ExxonMobil
George Smith	Sierra Club
Jim Smith	TCEQ
Steve Smith	Lyondell Equistar
Erik Snyder	EPA Region VI
Randall N. Stowe	Dow
George Talbert	TARC
T.W. Tesche	Alpine Geophysics
Don Thompson	TCEQ Houston Regional Office
Alan Timme	Hunstman Corporation
Usha-Maria Turner	Texas Utilities
Lilly Wells	HGAC
Mike White	ExxonMobil
Shelley Whitworth	HGAC
John Wilson	GHASP
Jim Yohn	BP America, Inc.
Steve Ziman	Chevron

2.6 Relation to Other Urban and Regional Modeling Protocols

This protocol revision (December, 2003) includes descriptions of the COMA plans: conceptual model development and episode selection, meteorological modeling, emissions inventory development, model performance evaluation, future case modeling, and procedures for documenting and archiving model results. The work described in this protocol is a continuation of work performed for the December, 2002 SIP revision, with a significantly expanded scope, and will make greater use of the rich ambient data sets and scientific analyses which resulted from the TexAQS 2000.

This protocol does not specifically address modeling which will be conducted for the Beaumont-Port Arthur (BPA) area, but there is much commonality between the modeling to be conducted in the two areas. Because of the proximity of the two areas to each other, much of the meteorological and emissions modeling conducted for one region will apply to the other. Any modeling to be conducted for the BPA area will be preceded by the development of a separate protocol.

3 Conceptual Model of Ozone Formation in the Houston/Galveston Airshed

A conceptual model is important to the modeling process because it incorporates the major factors affecting ozone formation and their interrelationships. The modeling staff has developed a conceptual model for HGB to aid in selecting episodes and in understanding the modeling in relation to the major factors involved. The conceptual model is fairly generic in that it focuses on patterns manifested over long periods of time, but does not characterize ozone formation on any particular day. The conceptual model is provided in Appendix A.

Additional insight can be gained through developing conceptual models specific to ozone episodes. The modeling staff contracted for the development of a detailed day-by-day conceptual model for the entire TexAQS 2000 period, including the August 18-September 6, 2000 episode. The primary purpose of these day-specific conceptual models is to provide bases for qualitative model performance evaluation, i.e. Does the model replicate features described in the daily conceptual model such as sea breeze timing, rapid ozone development, etc.?

Daily conceptual models for the entire TexAQS 2000 period are provided in Appendix B **Meteorological and Ozone Characteristics in the Houston Area from August 23 through September 1, 2000**, which is a contractor report and can be located at:

http://www.tnrcc.state.tx.us/air/aqp/airquality_contracts.html#section3

Despite the report's title, the report considers the entire period from August 15 to September 15, 2000 (although somewhat more attention is focused on the original August 23-September 6 episode period).

4 Domain and Database Issues

4.1 Air Quality Data and Meteorological Databases

4.1.1 Surface Measurements

The TCEQ routinely measures meteorological parameters and ozone and NO_x concentrations at continuous monitoring sites in Harris, Galveston, Montgomery, Brazoria, Jefferson, and Orange counties. The City of Houston measures various meteorological parameters and ozone and NO_x concentrations at an additional seven sites. During the year 2000, the TCEQ purchased ozone, NO_x, and meteorological data from the Houston Regional Monitoring Corporation and the Southeast Texas Regional Planning Commission, and these two organizations are donating the remainder of their data to the data set for TexAQS 2000.

4.1.2 Upper Air Measurements

Data from the six radar wind profilers, three rawinsonde launch sites, and multiple surface measuring stations were assembled into a single quality assured data set to support data analysis and MM5 meteorological modeling. This modeling generated the wind fields to support photochemical grid modeling of the August 23 through September 1, 2000 episode during the TexAQS 2000 period. This rich data set was used to assess performance of the MM5 modeling being conducted by Texas A&M. In addition, Texas A&M was responsible for formatting the profiler data for observational data assimilation.

Several nudging issues were investigated in different model runs and reported in previous SIP documents. During the extended TexAQS episode decisions on where and how much nudging should take place and how much data to reserve for performance evaluation will be made by TCEQ, ENVIRON, and ATMET modeling staff.

Upper air data was used in Phase 1 modeling to help determine the appropriate range for available soil moisture. The vertical profiles of observed data were also used to evaluate the physics options selected in MM5. This approach was followed initially for the COMA, but was supplanted (for part of the extended episode) by a new approach using data from the Geostationary Operational Environmental Satellite (GOES). As is discussed in more detail later in this Protocol, GOES data is being used to provide surface temperature observations and solar insolation for assimilation into MM5. With this data available soil moisture can be adjusted to minimize predicted and observed temperature tendencies.

The same nudging file used in the December 2000 SIP revision is being used for the core TexAQS period (August 25 - September 1). Upper air profiler wind data analyzed by Dr. Allen White of National Oceanic and Atmospheric Administration (NOAA) is being used for the additional portions of the extended TexAQS period (August 18 - 24 and September 2 - 6).

4.1.3 Aircraft Measurements

Data from the four aircraft making chemical measurements during TexAQS 2000 are being used in both modeling and non-modeling analyses. The data from the aircraft are providing some of the most important new insights into the cause of unusually high ozone in the greater Houston area. TCEQ has used the meteorological data collected by aircraft as part of its evaluation of both MM5 and CAMx performance, and will continue to expand its use of this data as the modeling progresses.

The chemical data collected by the various aircraft have greatly improved understanding of the causes of ozone formation in the HGB area. Analyses by NOAA and Brookhaven National Labs indicated ozone yields much higher than those seen in other cities, and linked these yields to the industries in the Houston Ship Channel as well as with smaller, more isolated petrochemical complexes in the area.

Further analyses of the aircraft data indicated that the modeling inventories are severely deficient in VOC's, especially light olefins such as ethene (ethylene) and propene (propylene). In modeling conducted for Phase 1 of the MCR, it was necessary to apply a top-down adjustment to the emissions of HRVOCs to achieve reasonable model performance. Since the completion of Phase I, a number of analyses of ambient concentrations vs. reported emissions of HRVOCs have been conducted, and additional work in this area is ongoing.

These studies, as might be expected, have yielded a range of adjustment factors which could be applied to emissions. The adjustment previously used in Phase 1 lies within this range, and in fact none of the studies conducted to date have been conclusively shown to provide superior adjustments.

While the major focus has been on HRVOC emissions, TCEQ has also studied other hydrocarbons, and we are sponsoring additional studies to assess the accuracy of reported inventories of these compounds. TCEQ staff have compared observed ratios of various hydrocarbons to NO_X to reported emission rates of these pollutants and have developed preliminary adjustments for several classes of compounds. Additional current and planned research should provide further refinement of these adjustments.

4.1.4 Satellite Observations

The TCEQ has recently begun incorporating data measured from satellite observing platforms into its modeling process. As in the Phase 1 MCR modeling, we are using solar radiation derived from observations made by the Geostationary Operational Environmental Satellite (GOES) to estimate Photosynthetically-Active Radiation (PAR) for input into the biogenic emissions model.

For the current modeling application, the TCEQ supported a version of MM5 which incorporates GOES solar insolation directly into the model and nudges surface temperature tendencies (derivative of temperature with time) to surface temperature tendencies obtained from GOES data. At this writing, this new version of MM5 has been run for the original TexAQS episode (August 22-September 1, 2000) and the results show significant improvement over the modeling conducted for Phase 1. Thus, this version of MM5 - with other options consistent with previous work - will be used for the core TexAQS period. If the remaining episode days can be processed in time for inclusion into the current SIP revision, we will attempt to model the entire episode using the new GOES-based MM5.

The TCEQ is currently sponsoring a project to derive photolysis rates from GOES data for incorporation into CAMx. If successful, this effort will provide more accurate estimates of a critical modeling parameter and should improve the model’s ability to simulate what actually happened on a given day.

4.1.5 Enhanced VOC monitoring in industry-sponsored monitoring networks

At TCEQ’s request, the industry-sponsored monitoring networks in Houston, Texas City, and Brazoria County have greatly enhanced their VOC monitoring capabilities by adding new sites and new monitoring equipment. The data collected at these sites and by this equipment is being provided to TCEQ, where it is being used to ascertain which VOCs are most important in generating ozone, and to what extent VOC emissions may be under-reported at industrial facilities. Three new PAMS automated gas chromatographs (auto-GCs) have been installed in the Houston area. One new PAMS auto-GC was installed in Texas City and three auto-GCs were installed in Brazoria County downwind of the Sweeney, Freeport, and Chocolate Bayou industrial complexes. Each auto-GC site includes ozone, NO_X, and meteorological monitoring. These data are also being provided to TCEQ.

Preliminary analysis of this data has begun, and TCEQ analysts are actively involved with the interpretation of this data as it arrives. If results obtained through analysis of the new data are available in time, they may be utilized in the current modeling application. If not, the data will be extremely valuable in modeling for future SIP revisions.

4.2 Modeling Episode Selection Process

4.2.1 Overview

TCEQ staff and the photochemical modeling Technical Committee (TC) meet periodically to review the development of the Conceptual Model for Ozone in the HGB area, evaluating episode selection criteria, and screening appropriate episodes for several years. As a result of discussions and interaction with the TC, the TCEQ staff reviewed recent ozone episodes between 1998-2000 and developed recommendations based upon the conceptual analysis discussed in Appendix A, the HGB Conceptual Model.

The committee recommended evaluating ozone episodes based upon their ability to represent the most frequent and typical patterns associated with current high ozone in the HGB. Therefore, the review was limited to the most recent three years of complete data (at that time, 1998-2000) to reflect the EPA procedures for determining the design value. The committee also placed special emphasis on using the results of the TexAQS 2000 study since it provides a very strong data base on which to evaluate hypotheses and validate modeling results.

Candidate episodes from the 1998-2000 period were selected in order to reflect the current design value and the current emissions. Ozone episodes in the HGB area occur most frequently in the August - September period. In fact, the design value for approximately 66% of the area monitors was established during the August - September 1998-2000 period, and the then-current design value for the entire HGB area (199 ppb at Clinton Drive, September 23, 1999) was also established during this period. Therefore, TCEQ decided to select episodes from this time period for Phase 1 modeling.

4.2.2 EPA Guidance for Episode Selection

EPA guidance for episode selection recommends analyzing the morning wind directions to determine the relative frequency of wind patterns and calms associated with ozone formation. (Guideline for Regulatory Application of the Urban Airshed Model, EPA-450/4-91-013, July 1991) The Guidance also recommends selecting for modeling one of the top three events from each wind direction associated with high ozone concentrations. During the COAST study, the TCEQ recognized the importance of the morning/afternoon land/sea breeze flow reversal and therefore enhanced the EPA method by including 4 hours of afternoon winds in the analysis. Review of the wind patterns during high ozone events using the improved TCEQ morning/afternoon method indicated that the wind directions change dramatically during the day as a result of the sea breeze flow reversal.

Recent analysis suggests a similar but more dynamic picture. Although it is still true that on high ozone days the early morning winds tend to come from the northwest and afternoon winds tend to come from the southeast, the picture is much more complex than originally thought. First, high ozone events occur when vertical mixing is limited and the winds are generally light and variable in direction. During ozone episodes, the average wind vectors are light and northwesterly in the early morning hours and veer clockwise through all the compass directions during the 24 hour day. The veering wind pattern tends to bring local emissions back over the city, resulting in high ozone concentrations.

In contrast, on low ozone days winds are stronger and do not veer continuously throughout the day. The wind direction is relatively persistent, suggesting ventilation and relatively low ozone. Given this clear distinction between high and low ozone events, it is important to show that the ozone episodes selected for modeling illustrate the typical patterns associated with ozone events.

However, detailed analysis of individual ozone events shows that day-to-day variations in wind direction do not always match the average pattern. The key issue in Houston ozone seems to be that regardless of direction, the winds during episodes are relatively light. Stronger, more persistent winds tend to dilute the ozone and blow it out of the HGB area.

4.2.3 TCEQ Staff Recommendations for HGB Ozone Episode Selection

Ozone episodes selected for modeling should represent the most frequent, typical, and representative patterns associated with high ozone in the HGB area. TCEQ staff recommended considerations are listed below.

- Ozone episodes occur most frequently during the August-September time period, when the design values at most of the area’s monitors are established. Therefore, this is the best time period from which to select additional episodes to model.
- Recent episodes are preferred over older episodes because recent episodes better represent the current emissions inventory, including mobile and point configurations.
- Well-monitored episodes (with more meteorology, VOC, and NO_x data) are preferred over data poor episodes. Additional data allow for more thorough model evaluation and provide information necessary to understand the processes leading to high ozone.
- The land/sea breeze flow reversal is associated with some of the worst exceedances seen in the area. At least one episode selected for this area should include this type of wind pattern. However, since wind directions vary from day to day, episode selection for the Houston area should include a variety of resultant wind directions. As discussed in the Conceptual Model (Appendix A) the most current thinking about the dynamics of ozone formation in the HGB area subsumes the land/sea breeze as part of a wind pattern in which the wind direction veers clockwise throughout the 24-hour diurnal cycle.

- Ozone episodes should be selected with monitored values within ±10 ppb of the design value to represent the magnitude of ozone that must be controlled. Furthermore, episodes should have high ozone in the geographical areas where high values typically occur. Such selection will allow testing control strategies with representative ozone concentrations in the areas where the strategies must work.
- Multi-day episodes are the most efficient way to model both long range transport of ozone as well as the accumulation of local emissions that are associated with ozone in the HGB area. Multi-day episodes also allow the possibility of testing model responsiveness. If the model can reproduce both the high and low ozone days within a period, there will be more confidence in the model and its ability to replicate events.
- High ozone occurs frequently in several different areas, but the highest design values occur in the east, near the ship channel. Episodes should represent the most frequent geographical patterns of exceedances as well as the hot spot areas reflected in the current HGB design value analysis.

4.2.4 Selection of Episodes for the COMA

Episodes occurring during the August - September period from the most recent 1998-2000 design value analysis period were screened, and three episodes were initially selected based upon a global assessment of the considerations listed above rather than any single acceptance or rejection criteria. The Technical Committee selected the August 2000 episode because it was a multi-day episode that occurred during an exceptionally well monitored period and represented typical ozone conditions. The two additional 1998 episodes were recommended at that time to address a broader range of conditions and to include days where additional monitors registered ozone peaks in other areas near their respective design values.

- August 25-September 1, 2000
- August 1-5, 1998
- August 26-30, 1998

The August 25-September 1, 2000 TexAQS 2000 Episode

The August 25-September 1, 2000 episode was selected because it had numerous exceedances in both Houston and Beaumont, so the episode was useful for both nonattainment areas. Six exceedance days (Ozone >124 ppb) occurred in the Houston area during the 8-day period, including two days with multiple-exceedances and a period of apparently low ozone in the middle. The low ozone days were initially selected to test the model's ability to respond to increases and decreases in daily ozone.

The episode includes 5 days with the veering winds typically associated with flow reversal and high ozone (see the HGB Conceptual Model provided as Appendix A). August 25 has light easterly winds resulting in maximum ozone at Crawford in the center of the Houston area. August 26 has southeasterly winds carrying the maximum ozone out of Houston to Conroe. Initially, August 27 and 28 appeared to be two low ozone days, with stronger southeasterly sea breeze winds resulting in substantially lower ozone in the HGB area and transporting the diluted urban plume toward Conroe. August 29, 30, and 31 have light westerly morning winds followed by weaker afternoon sea breeze winds which position the ozone pool on the east side of the city at Mt. Belview, La Porte and Deer Park. September 1 has a relatively persistent westerly land breeze, which carries the maximum ozone to the Baytown monitor and points further east.

4.2.5 The TexAQS 2000 Extended Episode

Subsequent to selecting the three episodes, extensive study of the TexAQS 2000 data has shown that the wind directions and exceedance areas that led to including the two 1998 episodes had also occurred during the TexAQS period. Also, the aircraft measurements during TexAQS provided some very important data and insight that was not available during the 1998 episodes. Finally, the exceedances that occurred in the 1998 episodes were much higher than the current design value, so those old episodes did not properly represent the current situation.

Analysis of the TexAQS aircraft data showed ozone plumes and areas of high ozone not reflected in data from the surface monitoring network. Therefore, high ozone was occurring on additional days and in areas that the preliminary analysis had not been previously associated with ozone formation. Since the TexAQS study included a larger and substantially improved database (including airborne measurements of NOx, VOC and other compounds, profiler measurements of mixing height and winds in the boundary layer as well as comprehensive LIDAR and chemistry measurements at the La Porte site and the Williams tower), TCEQ staff decided that the dynamics associated with the two 1998 episodes would be better represented by extending the TexAQS modeling window to include the "missing" sites and wind directions.

Table 3 shows some of the statistics for the extended TexAQS modeling window. Red text marks the dates where the aircraft or the Williams Tower measured higher ozone than the surface stations. Underlined text indicates where aircraft measured ozone greater than 125 ppb when the surface monitoring network suggested that no exceedances had occurred on that day.

The extended TexAQS episode also now includes 13 exceedance days measured by the surface monitoring network, numerous days when aircraft measurements were higher than surface measurements, and four additional high ozone days identified by aircraft and the Williams Tower data. The extended episode now includes three days with multiple exceedances at 9 to 12 surface monitors, and other periods with ozone occurring at from 1 to 7 surface sites. This variety will provide an excellent test of model dynamics.

The additional days added by the extended episode provide some valuable additional information, while picking up ozone in locations and wind directions not previously represented in the shorter episode used in Phase 1.

- The beginning of the extended episode (Aug 19-22) provides some interesting ramp up days with surface and aircraft exceedances, and adds another classic flow reversal episode (Aug 21) with numerous exceedances that appear sequentially at different sites.

- The additional days at the end of the period (Sept 2- 6) pick up coastal exceedances driven by the land breeze (Texas City, Galveston, Clute) as well as an interesting transport event with two extended ozone plumes measured by aircraft and supported by a single surface exceedance at Croquet on Sept 6th.

Finally, but perhaps most important, since the extended TexAQS episode occurred entirely during the TexAQS 2000 special study period, it allows access to all the special surface and airborne meteorological and air quality measurements and scientific analysis accomplished during the period.

The August 16-September 6, 2000 ozone episode occurs during the peak of the ozone season and includes a full suite of daily wind directions which is indicative of a full synoptic cycle. It also includes days with persistent land breezes and days with stagnation/flow reversal, as well as 13 one-hour exceedances and 14 eight hour ozone exceedances. For these reasons, TCEQ staff considers the episode to be fully representative of typical ozone patterns in the Houston area.

Table 3: Summary of the August 19 - September 6, 2000 Extended Episode

Episode Day	Measured Sfc Max Ozone	Peak Station	# Sfc Stations Exceeding	Aircraft Measured Ozone	Flow Reversal?
August 19	146 ppb	Mt Bellview	1	168 ppb	Yes
August 20	113ppb	Mt Bellview	0	130 ppb	Yes
August 21	159 ppb	Hou Reg Ofc	9	210 ppb	Yes*
August 22	107 ppb	Aldine	0	80 ppb	Yes
August 23	101 ppb	Bayland Park	0	149 ppb**	Yes
August 24	120 ppb	La Porte	0	128 ppb	Yes
August 25	194 ppb	Crawford	12	233 ppb	Yes*
August 26	140 ppb	Conroe	1	152 ppb	Yes
August 27***	87 ppb	Conroe	0	115 ppb	Sea Breeze
August 28***	112 ppb	Conroe	0	140 ppb	Sea Breeze
August 29	146 ppb	Mt Belview	3	211 ppb	Yes
August 30	200 ppb	La Porte	7	220 ppb	Yes
August 31	175 ppb	La Porte	10	194 ppb	Yes
September 1	163 ppb	E Baytown	2	210 ppb	Land Breeze
Sept 2	125 ppb	Deer Park	1	---	Sea Breeze
Sept 3	127 ppb	E Baytown	1	153 ppb	Sea Breeze
Sept 4	164 ppb	Texas City	2	132 ppb	Yes
Sept 5	185 ppb	Galveston	3	239 ppb	Land Breeze
Sept 6	156 ppb	Croquet	1	160 ppb	Land Breeze
Totals	13 Exc Days	---	53 Exc Sites	17 Exc Days	

* Classic Flow Reversal Case with numerous monitors showing sequential exceedances.

** High Ozone measured at Williams Tower.

*** Days previously thought to have low ozone and no exceedances

4.3 Selection of Air Quality Model

For air quality models to be successfully used as technical support for a regulatory initiative, they must be physically sound. The model performance evaluation described in Section 6 is designed to determine whether the model is a valid tool for identifying potential control strategies. In a regulatory environment it is crucial that oversight groups (e.g., EPA), the regulated community, and the interested public also be convinced of the suitability of the model.

To ensure that a modeling study is defensible, the model must be scientifically appropriate for the intended application and be freely accessible to all stakeholders. The following three simple prerequisites were set for selecting the photochemical grid model to be used in previous rounds of the HGB attainment demonstration.

- Must have a reasonably current, peer-reviewed, scientific formulation.
- Must be available at no or low cost to stakeholders.
- Must not require the reformatting of available model inputs from earlier rounds of the study.

The only model to meet all three of these criteria is CAMx. The model is based on well-established treatments of advection, diffusion, deposition, and chemistry. Another important feature is that NO_X emissions from large point sources can be treated with the plume-in-grid submodel that helps avoid the artificial diffusion that occurs when point source emissions are dumped into a grid volume. The model software and the CAMx user's guide are publicly available at http://www.camx.com.

When CAMx was selected as the primary model for Texas, an important consideration was that the required CAMx input files have the same structure as those of the Urban Airshed Model with Variable grid spacing (UAM-V). This stipulation allowed the previously generated COAST input files to easily be used in earlier modeling work. Since nearly all the modeling input files will need to be rebuilt for the current round of the modeling, the compatibility of file formats with UAM-V is less important than in previous modeling. However, TCEQ plans to continue using CAMx since the staff have many years of experience with CAMx, it is consistent with modeling being conducted in the near nonattainment areas, and it continues to offer state-of-the-science features and performance.

The TCEQ is also considering using EPA's Community Multiscale Air Quality (CMAQ) model in future work, but it would not be practical to transition to CMAQ for the COMA. The TCEQ has contracted for a detailed comparison of CMAQ and CAMx in Southeast Texas, and the results of this work will help guide TCEQ's decision of whether and when to adopt the CMAQ.

The current version of CAMx, Version 4.03, has a number of important features including:

Parallel Processing: **-** Multi-processor support is now fully included in CAMx. The approach is to use OpenMP compiler directives within the CAMx code. The chemistry and transport algorithms have been "parallelized" because they are the largest contributors to the CPU time.

IEH and CMC fast chemistry solvers for both CB-IV and SAPRC - Besides four versions of the Carbon Bond IV (CB-IV) chemical mechanism, CAMx provides the option to use the 1999 version of the SAPRC chemical mechanism. The SAPRC99 mechanism was added as an alternate mechanism because it is chemically up-to-date and has been tested extensively against environmental chamber data. Users can select between the SAPRC99 and the CB-IV chemistry at run time.

"Flexi-nesting" option - "Flexi-Nesting" allows nested grids to be introduced during a simulation (at a restart) and allows CAMx to interpolate any or all of the input data for a nested grid from parent grids;

Multiple "probing" tools for sensitivity analysis, including:

> **Process Analysis (PA) -** PA adds algorithms to the CAMx model that store the integrated rates of species changes due to individual chemical reactions and other sink and source processes. By integrating these rates over time and outputting them at hourly intervals, PA provides diagnostic outputs that can be used to explain model simulation in terms of chemical budgets, conversions of chemical species, and effects of transport and other sink and source terms. Process analysis can also improve model validation and ultimately can assist in the selection of precursor reduction strategies (Tonnesen, 2001).
>
> **Ozone Source Apportionment Technology (OSAT) -** OSAT provides a method for estimating the contributions of multiple source areas, categories, and pollutant types to ozone formation in a single model run. OSAT also includes a methodology for diagnosing the temporal relationships between ozone and emissions from groups of sources.
>
> **Anthropogenic Precursor Culpability Assessment (APCA) -**. APCA differs from OSAT in recognizing that certain emission groups are not controllable (e.g., biogenic emissions) and that apportioning ozone production to these groups does not provide information that is relevant to control strategies. To address this, in situation where OSAT would attribute ozone production to non-controllable (i.e., biogenic) emissions, APCA re-allocates that ozone production to the controllable portion of precursors that participated in ozone formation with the non-controllable precursor. In the case where biogenic emissions are the uncontrollable source category, APCA would only attribute ozone production to biogenic emissions when ozone formation is due to the interaction of biogenic VOC with biogenic NO_X. When ozone formation is due to biogenic VOC and anthropogenic NO_X under VOC-limited conditions (a situation in which OSAT would attribute ozone production to biogenic VOC), APCA re-directs that attribution to the anthropogenic NO_X precursors present.
>
> **Geographic Ozone Assessment Technology (GOAT) -** GOAT does not attempt to trace ozone production back to the source of the precursors, but rather ozone formation is tracked based on the geographic location where it occurred. Thus ozone

formation in a grid cell over “Area A” would be attributed to Area A even if the culpable emissions originated upwind in Area B. The disadvantage of GOAT is its simplistic assumption regarding the basis for ozone culpability. The advantages of GOAT are: (1) its freedom from assumptions about whether ozone formation is NO_X or VOC limited; and (2) that its results may be more directly comparable to other emerging source attribution methodologies (e.g., trajectories, tracers, and possibly the spatial analysis component of Process Analysis).

Direct Decoupled Method (DDM) - DDM provides an efficient and accurate methodology for calculating first order sensitivities between output concentrations and model input parameters;

There are many upgrades, bug fixes, and new features in CAMx 4.03. Details of these changes can be found in the Version 4.03 Release Notes, available on the CAMx website http://www.camx.com. TCEQ will continue to adapt the latest version of CAMx in this modeling project.

So far, TCEQ has run CAMx using the CB-IV chemistry, but the tools necessary to produce a SAPRC-speciated inventory have been assembled as well. The greater level of detail available through the SAPRC mechanism may prove valuable in simulating the complex photochemistry occurring in the HGB area, particularly in the industrial sectors. TCEQ plans to run side-by-side comparisons of CB-IV and SAPRC in the spring of 2004, and will consider using SAPRC in subsequent work if it compares favorably with CB-IV in the current application.

The CAMx implementation of chlorine chemistry was discussed by the SCC Modeling committee in 2002, but was not generally accepted due to limited peer review. TCEQ plans to test chlorine chemistry in the COMA analysis, but the results will only be presented as sensitivity analyses unless: 1). the inclusion of chlorine chemistry is seen to have significant control strategy implications, and 2). greater support for its use among the SCC members is forthcoming.

Several comparisons of CAMx with and without chlorine chemistry have been conducted by Environ and the University of Texas. Modeling results for the August and September 1993 episodes show that chlorine chemistry accelerates ozone formation by as much as 16 ppb/hour during the mid-morning hours near chlorine sources. However, the afternoon peak ozone concentration is only increased 1-3 ppb with the base case emission inventory. A combined chlorine and paraffin or olefin upset can increase the afternoon peak ozone concentration by as much as 6 ppb.

TCEQ is running CAMx v4 using the Piecewise Parabolic Method (PPM) advection solver. Since the relative extent of atmospheric transport and mixing which varies daily through the episode is not known *a priori*, the preliminary choice of advection solver is somewhat arbitrary. PPM may include less numerical diffusion than Smolarkiewicz, and it may be easier to accommodate future improvements to horizontal mixing with this advection solver in place.

4.4 Modeling Domain and Horizontal Grid Cell Size

Figure 3 shows the grid configuration for the COMA modeling. The CAMx modeling domain will consist of a 4 km × 4 km grid encompassing the Houston/Galveston and Beaumont/Port Arthur ozone nonattainment counties (light blue box), nested within a 12 km × 12 km grid covering the eastern part of Texas (green box). The outer 36 km × 36 km grid (blue box) was selected based on preliminary analyses using Hy-Split back trajectories, indicating that the domain as shown is sufficiently large to minimize the contributions of boundary conditions on the inner grid for the episode selected for the Phase 1 MCR modeling. Tracer simulations also indicated that the domain is sufficient for the COMA episode (August 18 - September 6, 2000). The outer grid may be changed if future analyses indicate that a different grid configuration is desirable.

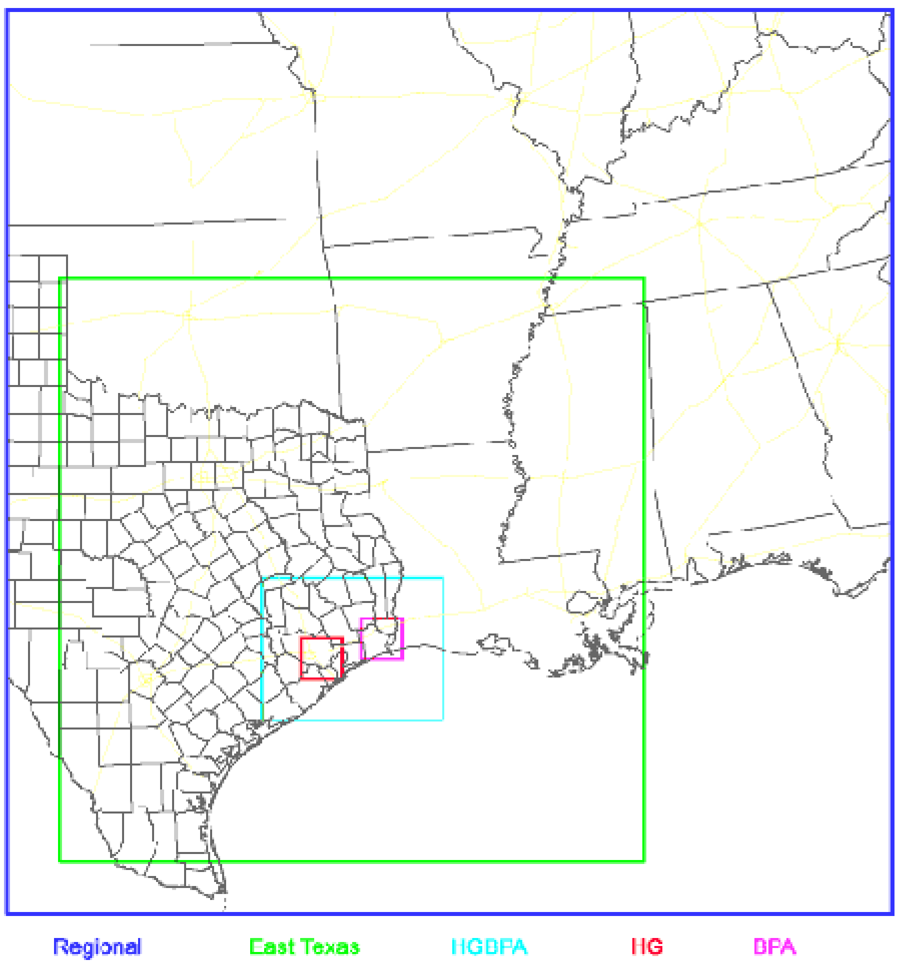

Figure 3: Grids selected for use in Combined Ozone Modeling Analysis

Also shown in Figure 3 are two superfine 1 × 1 km grids, one centered on the Houston Ship Channel (red box), and one covering the Beaumont-Port Arthur industrial area (purple box). Because of the computational expense required to model the superfine grids, we do not anticipate using these grids routinely. Sensitivity analyses were conducted to analyze the effects of these grids on the model's results, and to help determine for which runs and for which episode days we will use the superfine grids.

All grids are projected in a Lambert Conformal Projection (LCP) with origin at 100° W. and 40° N., and align with EPA's National Grid which was developed for nationwide modeling for haze and particulate matter. Choosing a grid system compatible with an existing large-scale grid system serves several functions, including providing ready-made regional inventory data which can be used directly, allowing TCEQ's modeling to be integrated into regional modeling projects, and promoting consistency among various regional and urban modeling applications in the central United States. Table 4 lists the grid dimensions for the CAMx domain and sub-grids for the COMA.

Table 4: CAMx modeling domain definition for the COMA

Grid Name	Grid Cell Size	Dimensions (grid cells)	Lower left-hand corner[1]	Upper right-hand corner[1]
Coarse Grid	36 × 36 km.	45 × 46	(-108, -1584)	(1512, 72)
Intermediate Grid	12 × 12 km.	89 × 89	(-12, -1488)	(1056, -420)
Fine Grid	4 × 4 km.	83 × 65	(356, -1228)	(688, -968)
HGB Superfine Grid	1 × 1 km.	74 × 74	(431, -1135)	(505,-1079)
BPA Superfine Grid	1 × 1 km.	74 × 74	(539,-1117)	(613, -1043)

[1]Grid corners are in kilometers (easting, northing) relative to grid origin at 100° W. and 40°N.

4.5 Number of Vertical Layers

The number of vertical layers is a compromise between including enough detail to accurately characterize the vertical layering of the atmosphere and managing the amount of time required to run the model. The TCEQ's Silicon Graphics modeling computer makes it feasible to employ many more vertical layers than have been used in past modeling exercises. Ideally, CAMx would be run with the same vertical layering as MM5; but since the latter uses sigma coordinates while CAMx uses standard height-above-ground-level, it is not possible to match the layers exactly.

The unique meteorology induced by the land/sea/bay effects and the unique mixture of industrial sources, which release pollutants across a wide range of elevations, indicate the need for many vertical layers, particularly near ground level. The Phase 1 MCR modeling employed 14 vertical layers. In this phase, the first eight CAMx layers closely matched the first eight MM5 layers. Layers 9 through 11 corresponded to two MM5 layers each, and layers 12 through 14 each corresponded with three MM5 layers. Detailed process analysis work conducted by Dr. Harvey Jeffries and his associates noted that modeled concentrations of ozone dropped sharply when the mixing layer extended into the 12th CAMx vertical layer. This phenomenon occured because the 12th CAMx layer was relatively thick compared with lower layers, since it corresponded to three MM5 layers. Combined with the fact that the MM5 layers grow progressively thicker with altitude, layer 12 had a thickness of approximately 750 meters, while the thickness of layer 11 was only about 285 meters.

For the COMA, the modeling staff designed a new 24-layer vertical structure in which the first 21 layers correspond with their MM5 counterparts. Three additional layers each correspond with two MM5 layers. This 24-layer structure is used within the 4 × 4 km. grid only (including the two superfine grids). A new 15-layer vertical structure is being used in the intermediate and coarse grids. Tables 5 and 6 below show the new vertical layer structure for the fine & coarse grids respectively. Note that the new structure extends to a height of 5836 meters above ground level (agl), compared with 4106 meters in the Phase 1 MCR modeling. The taller grid system helps to further insulate ground-level ozone concentrations from the top boundary conditions.

Table 5: CAMx Vertical Layer Structure for Fine (and Superfine) Grid

CAMx Layer	MM5 Layers	Top (m AGL)	Center (m AGL)	Thickness (m)
24	26, 27	5835.9	5367.0	937.0
23	24, 25	4898.0	4502.2	791.6
22	22, 23	4106.4	3739.9	733.0
21	21	3373.5	3199.9	347.2
20	20	3026.3	2858.3	335.9
19	19	2690.4	2528.3	324.3
18	18	2366.1	2234.7	262.8
17	17	2103.3	1975.2	256.2
16	16	1847.2	1722.2	256.3
15	15	1597.3	1475.3	249.9
14	14	1353.4	1281.6	243.9
13	13	1209.8	1139.0	143.6
12	12	1068.2	998.3	141.6
11	11	928.5	859.5	137.8
10	10	790.6	745.2	90.9
9	9	699.7	654.7	90.1
8	8	609.5	564.9	89.3
7	7	520.2	476.0	88.5
6	6	431.7	387.8	87.8
5	5	343.9	300.4	87.0
4	4	256.9	213.7	86.3
3	3	170.5	127.7	85.6
2	2	84.9	59.4	51.0
1	1	33.9	16.9	33.9

Note: AGL - Above ground level.

Table 6: CAMx Vertical Layer Structure for Intermediate & Coarse Grids

CAMx Layer	MM5 Layers	Top (m AGL)	Center (m AGL)	Thickness (m)
15	24, 25, 26, 27	5835.9	4970.9	1730.0
14	21, 22, 23	4105.9	3565.9	1080.0
13	18, 19, 20	3025.9	2564.5	922.9
12	15, 16, 17	2103.0	1728.1	749.8
11	13, 14	1353.2	1210.6	285.2
10	11, 12	1068.2	929.3	277.5
9	9, 10	790.6	700.0	181.0
8	8	609.5	564.9	89.3
7	7	520.2	476.0	88.5
6	6	431.7	387.8	87.8
5	5	343.9	300.4	87.0
4	4	256.9	213.7	86.3
3	3	170.5	127.7	85.6
2	2	84.9	59.4	51.0
1	1	33.9	16.9	33.9

Note: AGL - Above ground level.

4.6 Meteorological Data Fields for input into CAMx

The meteorological data fields required by the CAMx were developed for the Phase 1 episode by John Nielsen-Gammon of Texas A&M University, using the MM5 prognostic model in non-hydrostatic mode. Dr. Nielsen-Gammon is the Texas State Climatologist and was the official forecaster during the TexAQS 2000 campaign. Details of this work can be found in the December 13, 2000 SIP Appendicies/Atachments.

New meteorological fields for the expanded TexAQS episode (including the “core” period) were developed by Environ Corp and ATMET under contract to the Houston Advanced Research Center (HARC). In this modeling, important physics option remained consistent with the earlier work discussed in detail in Attachment 2 of the December 13, 2002 SIP Revision**.** MM5 radiation calculations, cumulus parameterization, explicit moisture calculations, and planetary boundary layer (PBL) algorithm were calculated using, respectively, the Rapid Radiative Transfer Model (RRTM), Grell, simple ice, and Medium Range Forecast (MRF) options of MM5. In contrast to Nielsen-Gammon’s work ATMET used the National Centers for Environmental Protection/Oregon State University/Air Force/Hydrologic Research Lab (NOAH) land surface model to predict available soil moisture. The NOAH land surface model is initialized by EDAS re-analysis fields. As discussed below in the section on “Planetary Boundary Layer Depth and Vertical Exchange Coefficients”, it was necessary to post-process the MM5-generated planetary boundary layer (PBL) depth to force reasonable agreement with radio profiler estimates.

Subsequent improvements to the MM5 modeling for the extended episode include using a version of

MM5 which assimilates GOES data for solar insolation (incoming radiation) and surface temperatures during the core TexAQS period of August 25th - September 1st. This version of data assimilation has made it unnecessary to post-process MM5-predicted PBL heights, hence is favored over previous meteorological characterizations. Unfortunately, the use of GOES assimilation was limited to the original (”core”) episode period and the remainder of the extended episode has not at this time been modeled using GOES. The TCEQ meteorological modeling staff are trying to acquire the software needed to prepare the remaining GOES data for input into MM5, and hope to be able to run GOES assimilation for the entire extended episode within the next several months.

Note that the meteorological modeling itself is not described in detail in this protocol. This section, as well as the next two, primarily discuss how meteorological data developed with MM5 are processed for input into CAMx. Since CAMx cannot directly ingest MM5 output, we have employed a program called MM5CAMx (available at no charge from Environ) to convert MM5 output data into a CAMx-ready format.

4.6.1 MM5 Modeling Domain

Figure 4 shows the MM5 modeling domain which remains unchanged from the Phase 1 modeling. The MM5 domain covers most of the North American continent, with an outer grid of 108 km × 108 km. The 36 km × 36 km, 12 km × 12 km, and 4 km × 4 km subdomains each correspond with CAMx grids as shown in Figure 2, except that the MM5 subdomains are slightly larger than their CAMx counterparts. Table 7 shows the grid configuration for the MM5 modeling.

Table 7: MM5 modeling domain definition for COMA

Grid	Dimensions (grid cells)	Lower left-hand corner[2]	Upper right-hand corner[2]
108 × 108 km.	52 x 42	(-2808, -2268)	(2808, 2268)
36 × 36 km.	54 x 54	(-324, -1728)	(1620, 216)
12 × 12 km.	99 x 99	(-72, -1548)	(1116, -360)
4 × 4 km.	149 x 134	(216, -1356)	(816,-816)

In addition to the above domains, in Phase 1 MM5 was run with a 1 x 1 km nested grid centered on the industrial Ship Channel area. The dimensions for this super-fine grid are as follows:

Table 7a: MM5 superfine grid

Grid	Dimensions (grid cells)	Lower left-hand corner[2]	Upper right-hand corner[2]
1 × 1 km.	96 x 80	(424,-1156)	(520,-1076)

After examination of the wind fields produced for the superfine grid, it was decided that these fields did not offer substantial improvement beyond merely interpolating the fine grid (4 × 4 km.) wind fields down to the 1 × 1 km. grid, so CAMx was never run using these very high-resolution wind fields. No

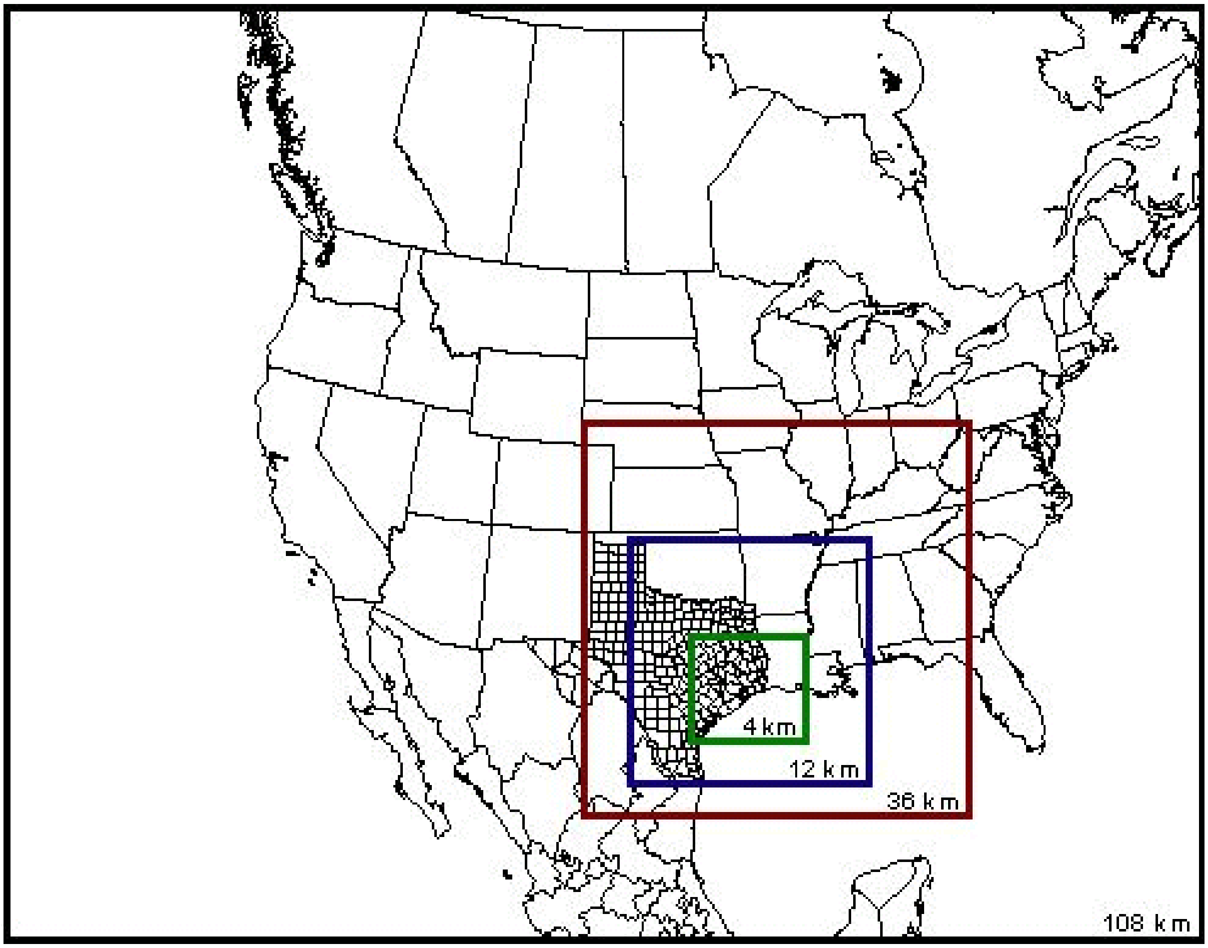

Figure 4: MM5 domain for COMA

further very high-resolution modeling using MM5 is planned at this time,

4.6.2 Wind Field Development

Wind fields are the most important product of the meteorological model for air quality modeling purposes. Winds mix pollutants and transport them across the domain. It is critical that the wind fields developed by the meteorological model represent the conditions seen in the actual ozone events, even though it is not reasonable to expect perfect replication. As is typical in air quality applications, upper air data was used to nudge the MM5 wind fields to more closely replicate the observations.

The present MM5 modeling incorporates three different types of nudging. First, MM5 wind fields are nudged towards the EDAS reanalysis fields on the two coarsest domains. Secondly, MM5 wind fields are nudged with observational profiler wind data on the four kilometer grid (and towards doppler lidar data at La Porte on August 25th). Finally, the use of GOES data, when available, allows for nudging of surface temperature tendencies (derivative with respect to time) which in turn allows for recalculation of available soil moisture. Each of these techniques influences the predicted MM5 wind field.

4.6.3 Planetary Boundary Layer Depth and Vertical Exchange Coefficients

PBL depth (sometimes referred to as mixing height) is a useful diagnostic for evaluating the potential impact of emissions and photochemical reactions on air quality. However, PBL depth itself is not an input parameter for CAMx. The MM5CAMx program is used to derive vertical exchange coefficients - which are input into CAMx - using a methodology developed by O'Brien (O'Brien, 1970). Vertical exchange coefficients (K_v)are used in the advection/diffusion equation to calculate mixing between adjacent vertical grid cells. In the present modeling, no adjustment to MM5-predicted PBL fields are made when running MM5CAMx during the core period of August 25th through September 1st. For episode days outside this period, we use a modified version of Environ's MM5CAMx program to scale the MM5-predicted PBL depths to agree more closely with the PBL depths derived from profiler data prior to calculating the vertical exchange coefficients.

4.6.4 Temperature

MM5 predicts hourly temperature values for each grid cell. The surface temperature is one of the variables which forces the growth of the boundary layer, and temperatures aloft are important for the stability of the atmosphere. Emissions of mobile sources and biogenic sources are temperature dependent, and in the photochemical model the temperatures affect the rates of chemical reaction rates. As mentioned earlier, GOES data was used to nudge surface derivatives, but not to nudge the surface temperatures directly.

4.6.5 Other Meteorological Parameters

Perturbation pressure and the water vapor mixing ratio are prognostic variables in MM5 which are also passed to a post processing algorithm which calculates five key photolysis rates. The photolysis rates also depend upon the solar zenith angle, altitude, and the spatially and temporally varying albedo, haze, and ozone column information provided by the Total Ozone Monitoring Spectrometric (TOMS) data.

It is necessary to post-process all the meteorological variables described above prior to input into CAMx. Environ (the developers of CAMx) have developed software to convert MM5 outputs to CAMx inputs. TCEQ has obtained the necessary software and has used it to convert the MM5 output data to CAMx-ready format.

4.6.6 Meteorological Model Performance Evaluation

TCEQ contracted with Environ for the development of a statistical package that allows Performance Evaluation (PE) to be conducted for meteorological modeling. The package is designed to interface with the MM5 model, which is the current model of choice. The package evaluates model performance for four meteorological factors: Wind Speed, Wind Direction, Temperature, and Humidity, and compares the model performance to benchmarks. Standardized statistics and graphs are developed for Bias, Root Mean Square Error, and Index of Agreement for each meteorological variable.

The PE package is currently being reviewed for implementation by the EPA, and has generated interest within the meteorological community. TCEQ intends to continue development and training, and is incorporating statistical methods into the performance evaluation of all current and future meteorological modeling.

4.7 Initial and Boundary Conditions

The modeling domain was selected to be sufficiently large to help minimize model sensitivity to boundary conditions. In addition, we begin the modeling three days prior to the first primary day of the episode to minimize the sensitivity to initial conditions. Default initial and boundary condition concentrations were used in Phase 1 and in preliminary modeling for the COMA. However, recent modeling analyses conducted in the Dallas/Fort Worth area by Environ showed an unexpectedly large sensitivity of ozone concentrations in that region to the lateral boundary conditions. Consequently, the default ("clean") boundary conditions were replaced by boundary conditions more representative of rural pollutant levels along the regional boundaries. To maintain consistency among modeling applications in Texas, we have adopted the DFW boundary conditions for use in the COMA. Sensitivity analyses have shown some improvement in HGB model performance using these somewhat higher concentrations, but the sensitivity to boundary conditions in the HGB region appears to be considerably less than that seen in the DFW modeling. We are working with several scientists from the Southern Oxidant Study (SOS) to assess the representativeness of these boundary conditions and if necessary to develop improved estimates of boundary concentrations for both the base case and for selected future years.

As discussed in the DFW modeling final report (available at http://www.tnrcc.state.tx.us/air/aqp/sipmod/dfwaq_techcom.html), the outer edge of the 36 Km. coarse grid was divided into three sections as shown in Figure 5 below (Note that the Dallas/Fort Worth coarse grid is identical to the one we are using for the HGB area). Boundary conditions for each of these segments were set to the values listed in Table 8. Initial concentrations were set equal to the values in the last column of the table.

Table 8: Boundary Conditions used in the COMA

Species	East/Northeastern Boundary Below 1700 m (ppb)	Western Boundary Below 1700 m (ppb)	Southern Boundary and Above 1700 m (ppb)
O3	40.0	40.0	40.0
NO	0.1	0.1	0.1
NO2	1.0	1.0	1.0
CO	200.0	200.0	100.0
PAR	14.9	14.9	14.9

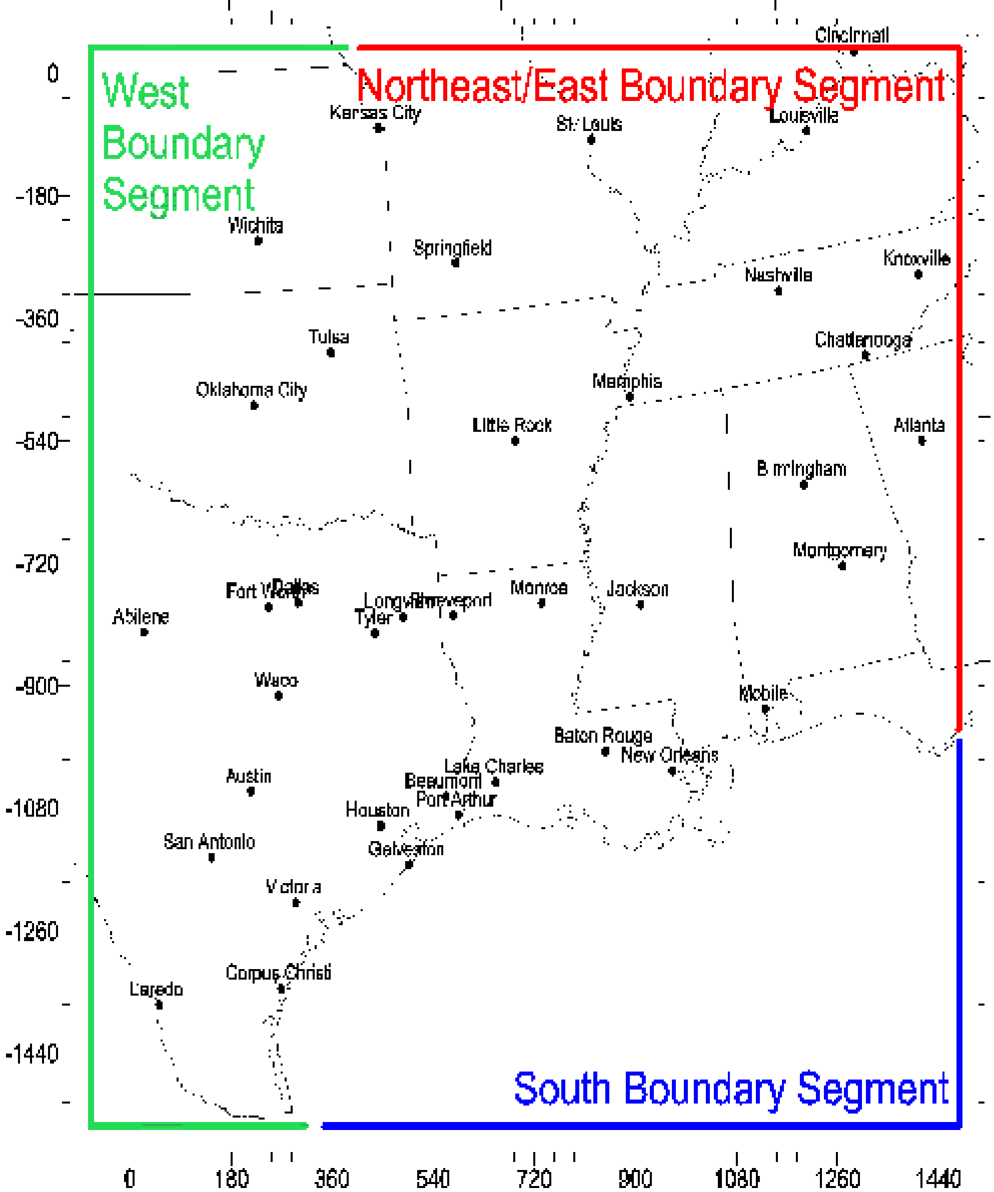

Figure 5: Segments used to define lateral boundary conditions.

HCHO	2.1	2.1	0.05
ETH	0.51	0.51	0.15
ALD2	0.555	0.555	0.05
TOL	0.18	0.18	0.0786
PAN	0.1	0.1	0.1
HNO2	0.001	0.001	0.001
HNO3	3.0	3.0	1.0
H2O2	3.0	3.0	1.0
OLE	0.3	0.3	0.056
XYL	0.0975	0.0975	0.0688
ISOP	3.6	0.1	0.001
MEOH	8.5	0.001	0.001
ETOH	1.1	0.001	0.001
Total NOx	1.1	1.1	1.1
Total VOC (ppbC)	50.5	22.3	9.3

4.8 Plume-in-Grid Modeling

CAMx has an option to model selected point sources with a PiG algorithm. PiG algorithms have historically been very computer resource intensive. However, the CAMx user's guide states that the PiG module within CAMx is considerably faster than PiG schemes in other models, and our experience with the new algorithm indicates that this is indeed the case. Additionally, with the computer resources now available, parsimonious PiG selection is no longer critical in terms of computer resource demands.

PiG sources were selected based on magnitude of NO_X emissions. As with Phase 1 of the MCR, over 300 PiG sources across the entire modeling domain, mostly large power plants, were selected. This number may increase or decrease as the modeling progresses.

5 Emissions Inventory

5.1 Base Case

The modeling emissions inventory (EI) is composed of point, area, on-road mobile, nonroad mobile, and biogenic emissions. The modeling inventory developed for the COMA contains data from a wide variety of sources, including the 1999 periodic inventory, the 2000 annual point source inventory, data from the ARPDB, data from the TexAQS 2000 Special Inventory, data from the TCEQ Region 12 Upset database, link-based on-road mobile source data, and data from several special studies, including a comprehensive inventory of plant species and biomass in East Texas. Day- and (in some cases) hour-specific inventories have been developed as appropriate, to account for temperature and activity variation. Future versions of this inventory will include emissions of chlorine from point and area sources.

5.2 Point Sources

The base case point source emission inventories are composed of information from several databases. For the Texas portion of the inventory, data from the point source database (PSDB) was used. A new modeling extract was queried from the PSDB, in order to capture any updates since the Phase 1 MCR modeling extract, but otherwise the basic point source inventory is essentially the same. As in Phase 1 modeling, the inventory is supplemented with hourly data from the Acid Rain Program database (ARPDB), data obtained during the TexAQS 2000 Special Inventory, and additional information from the TCEQ's database of upset/maintenance reports.

Emissions from both the PSDB and the Special Inventory contain large amounts of information about specific hydrocarbons emitted by each source; however, some sources report little or no speciation of their hydrocarbon emissions. In Phase 1 modeling, any source which reported less than 75% speciation was assigned either a Texas-specific SCC-average or an EPA default speciation profile. For sources reporting 75% or more speciation, the unspeciated emissions were assumed to have the same speciation as the reported emissions. This method is a major improvement over simply assigning default speciation based on SCCs, but still leaves some less-than-desirable results. Specifically, for any source whose emissions are less than 75% speciated, all reported speciation data is ignored. For COMA, we have developed a new process which retains all speciated hydrocarbon data reported to the PSDB, regardless of how completely each point's emissions were speciated. Also new for COMA speciation is the exclusion of non-VOC species, as defined by EPA, from all point-source speciation profiles. These procedures are described in "Speciation of Texas Point Source VOC Emissions for Ambient Air Quality Modeling", G. Cantu, TCEQ, October 2003.

The Louisiana Department of Environmental Quality (LDEQ) supplied to TCEQ modeling staff a copy of their 2000 point source emissions inventory in AFS format. The TCEQ modeling staff, with assistance and QA from LDEQ point source emissions staff, completed an AFS-to-ARPDB cross-reference list. This list links Louisiana Acid Rain boilers to their corresponding LDEQ stack identifiers. TCEQ modeling staff replaced LDEQ annual emission records in the AFS file with corresponding hourly ARPDB emissions for each hour of the episode.

For the states in the remainder of the modeling domain, beyond Texas and Louisiana, TCEQ used the same "regional" files generated for Phase 1 MCR modeling. Specifically, the TCEQ obtained point source emission records in the AIRS Facility Subsystem (AFS) format from Environ, Inc. This data had already been prepared for near-nonattainment modeling that Environ performed for several areas of Texas. TCEQ modeling staff reviewed the AFS file, removed Texas and Louisiana records from the file, and processed the remainder through EPS2x. TCEQ modeling staff created an AFS-to-ARPDB cross-reference list for the regional boilers larger than 750 MW capacity that are subject to EPA's Acid Rain Program. This cross-reference list links these boilers to their corresponding NEI/AFS stack identifiers. With this cross-reference file, the ozone-season daily emission records in the AFS file were replaced with corresponding hourly ARPDB emissions for each hour of the modeled episode.

TCEQ modeling staff has been in contact with the Minerals Management Service (MMS) over the last several years to monitor the status of the 2000 Gulf-Wide Emission Inventory (GWEI). As of this writing, the data have not been provided to TCEQ, so will not be used in the current round of modeling.

Therefore, the offshore point source emissions used for COMA will be the same as those developed during Phase 1 of the MCR. In Phase 1 of the MCR, the 2000 offshore EI was generated by growing the 1992 MMS offshore EI, in-place, by a factor that accounted for the growth in offshore production platforms, based on a previous MMS report. Based on the recommendation of MMS staff, we grew the entire point source offshore file by 44%, assuming that the ancillary stationary point source equipment would grow at the same rate as the number of offshore platforms.

The introduction of Mexican point sources was new to Phase 1 MCR modeling. TCEQ modeling staff converted the 1999 Big Bend Regional Aerosol and Visibility Observational (BRAVO) Study emissions inventory from IDA format to AFS format. This same Mexico emissions file was incorporated into COMA modeling runs. TCEQ modeling staff has completed a preliminary evaluation of the ERG July 2003 “1999 Mexico NEI” report and determined there were no significant differences in point source emissions. Additionally, the ERG data files have not been made available.

5.3 Adjusting the Point Source emissions based on ambient measurements

As was discussed extensively in the Technical Support Document of the December, 2002 SIP Revision, one conclusion of the TexAQS 2000 study is that observed concentrations of certain compounds, especially light olefins, are much larger than represented in the reported emissions inventories. In Phase 1 MCR modeling, using the reported emissions resulted in a severe under-prediction bias in modeled ozone concentrations, but when a set of highly-reactive VOCs (HRVOCs) were adjusted, model performance markedly improved.

The adjustment used in Phase 1 modeling consisted of creating a second point source emissions file in the standard AFS format used by EPS2x, containing all emission points for twenty-seven large HRVOC-emitting accounts in the eight-county nonattainment area. This file was used to provide the extra HRVOC emissions necessary to make each of the 27 facilities’ HRVOC emissions equal their individual NO_X emissions. The HRVOC-to-NO_X adjustment was based on the observation that airborne concentrations of light olefins measured aboard the Baylor University research aircraft frequently approximate concurrently measured concentrations of NO_Y when the aircraft passed through industrial plumes. Since the completion of Phase 1 modeling, several additional studies have been conducted comparing reported inventories to ambient measurements, both airborne and ground-based. These studies generally agree that emissions of HRVOCs are significantly under-reported. Additional studies are underway, and TCEQ plans to develop refined inventory adjustments in the near future.

The approach used in Phase 1 of the modeling is supported by at least one independent study conducted for the Houston Advanced Research Center by Environ (see http://www.harc.edu/harc/Projects/AirQuality/Projects/Status/Files/H6EDraftReport.pdf). This study used inverse modeling to assess various inventory components, and concluded that further modification of the inventory used in Phase 1 was not warranted under the then-current model formulation. For COMA, however, TCEQ is using a somewhat enhanced version of the adjustment used in Phase 1.

The Phase 1 approach has been enhanced in several ways. Most importantly, instead of adjusting all HRVOC species (which included a small adjustment of emissions of non-olefinic compounds), TCEQ has specifically targeted terminal olefins, since these are the compounds to which the aircraft instruments theoretically respond best. Second, instead of adjusting emissions at only a few selected facilities, TCEQ uses a broad-based adjustment which applies to all sources reporting emissions of more than 10 tons/year of terminal olefins. Third, the file used to boost emissions now contains explicit hydrocarbon species appropriate to each adjusted emission point, instead of the "generic" HRVOC used in Phase 1. Overall, these enhancements change the modeled reactivity slightly from Phase 1, but provide for much more flexibility in control strategy modeling.

The TCEQ plans to conduct additional studies comparing ambient concentrations of olefins to the inventory, and will work towards developing more targeted adjustments, especially now that several new automatic gas chromatographs (Auto-GCs) have been deployed in the industrial sectors of the HGB area. We also will study emissions of less-reactive VOCs to determine if and by how much these compounds are under-represented in the reported inventory. As discussed earlier in this Protocol, some preliminary assessments of emissions of these compounds have already been conducted, and we use the results of these assessments to develop sensitivity analyses of adjustments to the less-reactive VOC emissions. The results of these analyses, which should help to characterize the role of less-reactive VOCs in producing ozone in the Houston area, will be included in the documentation for the upcoming SIP revision.

5.4 Area and Nonroad Mobile Sources

Within the four-kilometer domain, area source emissions developed for the base case by projecting the 1999 periodic emissions inventory to 2000. Emissions from nonroad sources (except for ships, airplanes, and locomotives) were generated using the NONROAD 2002a model. For several categories, local equipment populations were estimated based on surveys: lawn and garden, recreational marine, and construction activity. Emissions for ships were estimated directly from a recent survey, and emissions for locomotives and aircraft were provided by the TCEQ Emissions Inventory staff. Special treatment was applied to shipping, with ship emissions treated as pseudo-stacks spaced along the major waterways within the Galveston Bay region (as described in the December 6, 2000 SIP revision).

Emissions for the remainder of Texas and for other states were obtained from Environ, who developed a 1999 inventory (based on the NEI) for modeling being conducted for the state's near-nonattainment areas. We recently received new statewide 2000 area source emission and will incorporate these into the current round of modeling as soon as practicable. We will also apply growth to the 1999 emissions used outside of Texas to produce a true 2000 base case for area and nonroad sources.

Spatial allocation for most categories employs new surrogates developed for the Phase 1 MCR modeling, including new spatial surrogates for shipping lanes.

Since we have not yet received the GWEI emissions estimates for area and nonroad sources, we will continue to use the same emissions as in Phase I MCR.

5.5 Mobile Sources

In August 2003, TTI provided MOBILE6.2-based updates to the HGB inventories for each day of the August 18-September 6 ozone episode for both the 2000 base case and the 2007 future case. The 2000 base case inventories were based on a maximum posted speed of 70 mph for various freeway segments, while the 2007 future case inventories were based on a maximum posted speed of 65 mph. As with previous development of on-road mobile source inventories for photochemical modeling purposes, TTI staff utilized travel demand model output for a specific episode year from the Houston Galveston Area Council (HGAC). For each roadway link in the eight-county HGB network, Vehicle Miles Traveled (VMT) and average speed estimates were developed for each hour of each episode day of interest. In order to distinguish between the differing traffic levels on the various episode days, TTI staff have developed adjustment factors based on in-use traffic survey data such as hourly traffic counts, VMT mix measurements, etc. MOBILE6.2 emission factor output in gram-per-mile by speed is coupled with the VMT per roadway link by hour to develop a complete on-road mobile source inventory of CO, NO_X, and VOC for the entire modeling episode in the eight-county HGB nonattainment area. All emissions are adjusted a final time to account for differences between the travel-demand model and the Highway Performance Monitoring System (HPMS). TTI is currently developing 2010 future case inventories for the August 18-September 6 episode in a similar fashion.

In July of 2002, TTI submitted MOBILE6 link-based inventory estimates for the 3-County Beaumont/Port Arthur (BPA) nonattainment area. These BPA inventories also covered the August 22-September 1 ozone episode for both the 2000 base case and the 2007 future case. In June of 2003, TTI submitted MOBILE6.2 updates to the 2000 BPA base case episode. At this time, TTI is developing 2010 future case inventories for this episode.

For the Texas counties within the modeling domain but outside the HGB and BPA nonattainment areas, HPMS-based VMT estimates were used by TTI to develop MOBILE6 county-wide emission inventories by roadway type for both 2000 and 2007. Due to the differing traffic profiles, inventories for each county were developed for Weekday (Monday-Thursday), Friday, Saturday, and Sunday day types. For preprocessing purposes, emissions from major roadways will be spatially allocated by appropriate roadway surrogates (e.g., interstates, state highways, arterials, etc.), while emissions from minor roadways and local streets will be allocated spatially by human population surrogates. Previously, 1999 MOBILE5-based inventories for these non-HGB counties were adjusted to MOBILE6 based on scaling factors from default runs with MOBILE5 and MOBILE6. These adjusted inventories were used in previous photochemical modeling efforts, but will be replaced by the MOBILE6 inventories developed by TTI once emissions preprocessing is completed.

In a similar fashion, emissions inventories for areas outside Texas but within the photochemical modeling domain utilized MOBILE5-based inventories which originated from EPA’s 1999 National Emissions Inventory (NEI). These inventories were also adjusted with MOBIE5-6 scaling factors. As time and resources permit, these inventories will be updated with Version 3 of the onroad NEI inventory, which was just released in December 2003.

TCEQ staff will continue preprocessing all of the MOBILE6.2 on-road mobile source inventory data using the EPS2x emissions preprocessor tool. PV-WAVE software will continue to be utilized as a quality assurance tool to ensure that the modeled on-road emission levels are properly distributed both spatially and temporally.

5.6 Biogenic Sources

Over the past five years, TCEQ has commissioned several studies for the purpose of improving the biogenic emissions estimates in Texas. These studies (Guenther et al., 2000; Yarwood et al., 1999; Yarwood et al., 2001; Wiedinmyer et al., 2000; Wiedinmyer et al., 2001) created a detailed vegetation map of Texas using field surveys and existing databases), and developed an operational version of the Global Biogenic Emissions Inventory System (GloBEIS) biogenic emissions model (Guenther et al., 1999).

Model **-** TCEQ is using the latest version of the GloBEIS, version 3 (Guenther et al., 2002; Yarwood et al., 2001; Yarwood et al., 2000; Guenther et al., 1999), to calculate biogenic emissions for this round of photochemical modeling. This version of GloBEIS includes several new features, including modules that vary the biogenic emissions according to changes in leaf area index, antecedent leaf temperatures, and drought, and an improved canopy energy balance model. TCEQ has performed some preliminary evaluations of new GloBEIS features and may include use of some of these in the current modeling evaluation, most likely as modeling sensitivities.

Vegetation data **-** The land use and vegetation database used for biogenics modeling is derived from three sources:

> TCEQ Texas vegetation database (Yarwood et al., 2000; Wiedinmyer et al., 2001). Based upon Texas Parks and Wildlife vegetation data, urban land use data from Braden, Collie, and Turner Consulting, agricultural statistics from the National Agricultural Statistics Survey, and field surveys carried out during 1999;
>
> BELD3 (Biogenic Emissions Landuse Data, version 3) (Kinnee et al., 1997). A vegetation database for the entire North American continent, prepared specifically for creating biogenic emissions inventories; and
>
> Mexican land use and vegetation database (Mendoza-Dominguez et al., 2000). Database created by researchers at the University of Monterrey and Georgia Tech.

The land use and vegetation database is gridded according to the Lambert Conformal Projection with reference origin at 40° N, 100° W. The data are available at 4-km resolution for a domain encompassing most of the states of Texas, Louisiana, Arkansas, Oklahoma, and Mississippi. The

TCEQ Texas vegetation database is being used for in-state biogenic inventories; the BELD3 database is being used for biogenic inventories in U.S. states outside of Texas, and the Mexican database is being used for Mexican biogenic inventories. If possible, the biogenic inventories may use enhanced vegetation data for the Houston area that is being assembled during 2002-2003 by a research project involving the Texas Forest Service, the U.S. Forest Service, the University of Houston, TCEQ, and the Houston Advanced Research Center. These data may be ready for use by January 2004 , and will encompass the eight-county ozone nonattainment area, with special emphasis on Harris County.

Temperature data - TCEQ is using temperature fields for biogenic emissions modeling created by spatially interpolating temperatures measured by NWS and other appropriate weather stations throughout southeast Texas. The density of measurement stations with high-quality temperature data in southeast Texas suggests that accurate temperature fields can be created by kriging (Reynolds et al., 2002). A recent paper by Vizuete et al. (2002) also suggests that kriging is the best interpolation method. TCEQ comparisons of MM5-derived temperature fields, kriging-derived fields, and independent ground observations of temperature suggest that kriging can generate fields that are at least as accurate as MM5 fields.

Photosynthetically-active solar radiation data (PAR) - TCEQ is using a new method for deriving PAR fields for biogenic emissions modeling. In the past, TCEQ used algorithms from the BEIS2 model to estimate solar radiation from cloud cover observed at ground-based weather stations. But this method can result in inaccuracies due to the uncertainties associated with interpolation, and to the somewhat subjective nature of cloud cover observations. Therefore, TCEQ is using PAR data derived from satellite measurements. These data are calculated by the University of Maryland and NOAA for the Global Energy and Water Cycle Experiment (GEWEX) Continent Scale International Project (GCIP). NOAA uses a modified version of the GEWEX surface radiation budget (SRB) algorithm (version 1.1) to calculate radiation flux fields from Geostationary Operational Environmental Satellite (GOES-8) data. In addition to the GOES-8 data, the algorithm uses ancillary information from the National Centers for Environmental Prediction Eta forecasting model to derive shortwave radiation fields at a regional scale. The algorithm's output is verified by comparison to ground-based solar radiation measurement stations. For further information about this method, see Pinker et al. (2003), Pinker and Laszlo (1992), and the GCIP/SRB web page at http://metosrv2.umd.edu/~srb/gcip/index.htm. TCEQ will be using a high resolution hourly database with spatial resolution of approximately 4 km^2. This database was developed specifically for TCEQ by Dr. Rachel Pinker at the University of Maryland (Pinker, 2002).

5.7 Future Year Emissions Development

After it has been determined that CAMx performs satisfactorily for the extended episode, the episode will be modeled with CAMx using a future-year inventory.

On-Road Mobile Sources

The 2007 on-road mobile source EI for the HGB nonattainment area was developed by TTI staff in a manner consistent with that described in Section 5.4. The main differences is that travel demand model output and MOBILE6 runs for 2007 were used instead of those for the base year. The 2007 travel demand model runs are based on best available projections of future population growth, demographic patterns, and roadway network changes. The MOBILE6 runs for 2007 utilize the same meteorological inputs as the base year (temperatures, humidity, etc.), but other inputs were modified as appropriate. Projecting into the future, it is expected that both the human and vehicle population in the HGB area will increase, thus causing an increase in daily VMT on the roadway network. However, typical turnover effects will yield a vehicle fleet more heavily populated with newer "cleaner" vehicles as opposed to older "dirtier" ones.

Point Sources

Phase 1 MCR modeling included essentially zero growth in the nonattainment counties, due to emission caps. Electric Generating Unit (EGU) growth in the attainment counties of East Texas was estimated via review of permit applications for large NOx sources that are expected to be operating by 2007. While we believe this is still valid, based on the general economic factors for the state, TCEQ has also committed to EPA to evaluate alternate growth procedures for non-EGUs in all of the attainment areas of our modeling domain. Analyses of several sources of economic growth data may be used to derive emissions "projection factors" for 2007 and 2010 for both the nonattainment and attainment areas of Texas.

The 2007 EGUs will be grown using the same procedure used in Phase 1 MCR. All newly operational, newly permitted, or proposed-operational EGUs between 2000 and 2007 within the state of Texas will be compiled into a separate database. Permit allowable emissions and stack parameters from permit files will be formatted into a single AFS file. This “new EGU” file will then be modeled as a separate stream of emissions using the emissions preprocessing system, so that it can be tracked and sensitivity analyses can be performed upon it. The 2010 EGU emissions will be estimated by growing the 2007 emissions with a 3-year growth factor derived from the growth evaluation study performed above.

As was done in Phase 1 MCR modeling, the banked emissions (ERCs and DERCs) expected to be used in/by 2007/2010 will be incorporated into the 2007/2010 projections.

MECT (Mass Emissions Cap and Trade) staff of the TCEQ maintain a database of sources that are subject to the HGB emissions cap. For Phase 1 MCR modeling, their database was compiled for the original “90% reductions” ESAD case. For the final Phase 1 MCR modeling, staff estimated the alternate ESAD (nominal 80% NOx reductions) capped emissions from the MECT database. Since Phase 1 MCR, MECT staff have recompiled their database for the new Chapter 117 “80% (alternate)” NOx ESADs. For COMA, modeling staff worked with MECT staff to model the most appropriate NOx emissions values for 2007 and 2010. Since the last phase of the HGB NOx Cap does not take effect until April 2008, a slightly different control level will be applied to estimate the 2007 NOx Cap as compared to the 2010 NOx Cap.

Discussions with State of Louisiana SIP staff indicated that no additional controls, beyond what TCEQ staff applied to the Baton Rouge area in Phase 1 MCR modeling, are expected in Louisiana by 2007. As with Texas, analyses of several sources of economic growth data may be used to derive "projection factors" for 2007 and 2010 Louisiana emissions.

TCEQ modeling staff are taking a new approach to modeling the Region (outside of Texas and Louisiana). In Phase 1 MCR modeling, TCEQ assumed no growth in the Region and only federal NOx SIP Call controls. EPA has performed some modeling in the last three years with a future case inventory in order to determine the impact of the federal Heavy-Duty Diesel (HDD) engine rules that are to be in place before 2007. TCEQ modelers obtained these inventory files from Environ International Corp., who processed the original EPA files and used them in 8-hour ozone modeling analyses for northeast Texas. This 2007 HDD inventory takes into account growth and controls, such as the NOx SIP Call, for all areas outside of Texas.

For our 2010 8-hour modeling, we intend to use another EPA-developed modeling inventory. The Clear Skies Act (CSA) modeling EPA completed within the last year, in order to determine the impact of reductions associated with the Clear Skies Act. The CSA, formerly the Clear Skies Initiative, is a potential replacement for the NOx SIP Call. EPA's CSA modeling was performed for the years 2010 and 2020. TCEQ has obtained these files from EPA, which are in MODELS-3 (CMAQ) IDA format, and has converted them for use with our emissions preprocessing system.

For COMA, offshore point sources emissions will remain at estimated 2000 levels for 2007 and 2010. Discussions with MMS staff have concluded that while there is likely to be substantial growth between 2000 and 2010, this growth will occur much farther offshore (50 to 100 miles) from Texas and Louisiana than in previous years. So it would be inappropriate to grow the existing sources in the traditional manner. For future modeling, we will work with the MMS to develop a suitable method for allocating future growth after we have obtained the GWEI.

Area and Nonroad Mobile Sources

In Texas, area source inventories will be grown to 2007 and 2010 using appropriate growth factors. Some categories may be projected instead using county-level economic and population growth projections from the Texas Comptroller's Office or special studies for categories like Oil and Gas production. Growth factors from EGAS will be used in cases where more current or locally-generated growth factors are not available. Controls will be applied as where appropriate.

Outside of Texas, for the 2007 future year, we plan to use the inventory developed for EPA for modeling HDDV rules. For 2010, we plan to use EPA's Clear Skies inventory.

For nonroad categories, the NONROAD model will be used to project emissions into the future since it accounts for both growth and federal controls on nonroad sources.

Biogenic Sources

Biogenic emissions are assumed to remain unchanged in the future, although urban development does modify the amount, location, and type of vegetation over time. TCEQ plans to investigate the use of projected land-use data to estimate attainment-year biogenic emissions in future modeling applications.

5.8 Emissions Processing

The TCEQ has at its disposal several software packages for processing anthropogenic emissions, including SMOKE, EPS-2, Fast-EPS, and EMS-95. In the Phase 1 MCR modeling, as well as in the COMA, TCEQ used a new version of EPS-2 known as EPS2x. This software is available from Environ, Inc. and executes much faster than the original EPS-2. In addition, it incorporates a new feature allowing modification of the model-ready emissions files at the county level. This feature means sensitivity runs can be conducted without re-running much of the EPS2x code. The EPS-2 family of emissions processors has several advantages over other systems including excellent reporting capabilities, stability, and ease-of-use. In addition, TCEQ staff are intimately familiar with the software and have developed numerous scripts and programs to interface with it. TCEQ processed all inventory components using EPS2x except for biogenics. Note that except for some enhancements, EPS2x is functionally equivalent to other versions of EPS-2, so it is expected to produce identical model-ready files in most cases. While TCEQ does not have sufficient resources to make a head-to-head comparison of EPS2x with the emissions processor used in previous modeling, the modeling staff is exercising due diligence to ensure that no errors have been introduced into this formulation of the emissions processor.

The TCEQ recently contracted with Environ to implement several enhancements into EPS2x. The new suite includes updates to several processing modules. Those updates include: expanded data fields and arrays, the ability to model hourly specific adjustments to mobile sources, flexible level-of-detail reporting schemes and alternative data reporting summaries. Some of these tools are currently in use and others will be used in future emissions processing as time allows for further testing and QA of results.

Biogenic emissions are being processed using the GloBEIS processing system used in the 2000 and 2002 SIP revisions, (or a new version if one becomes available).

5.9 Modeling Inventory Performance Evaluation

Aside from performing extensive quality assurance of the modeling inventory while it is being developed, TCEQ is performing several comparisons between the modeling inventory and ambient measurements. Because direct comparisons between emission rates and ambient air measurements are not meaningful, these comparisons are usually relative comparisons among measured compounds. For example, the ethane/NO_X ratio calculated from the emissions inventory can be compared to the same ratio calculated from ambient measurements. These comparisons should be made with care to ensure that the observed ambient air was actually influenced by the source of interest, and that at least one of the species in the ratio is relatively well quantified. If used judiciously, however, these comparisons can give some insight into possible shortcomings of the modeling inventory.

To date, numerous studies have compared ambient measurements with the reported emissions inventory. Researchers at several institutions including the TCEQ have compared aircraft measurements with the reported inventory and have concluded that the reported emissions of certain highly-reactive hydrocarbons, particularly light olefins, were significantly under-reported in the inventory. In addition, TCEQ staff have compared the reported emissions of light olefins with measurements made at automatic gas chromatographs (auto-GCs) in the area, and have reached similar conclusions. To date, no unified picture has emerged other than the need to significantly increase modeled emissions of light olefins. Additional studies are currently underway or are planned, including a tracer study at a major refinery in the HGB area. Results of these studies will help provide improved adjustments both for light olefins and as well as for other compounds in future modeling applications. For the meantime an enhanced version of the Phase 1 adjustment methodology is being used in the COMA.

6 Model Performance Issues

6.1 Quality Assurance Testing of Inputs

At each step prior to conducting base case simulations, the input fields will be reviewed for consistency and obvious errors. Graphical and statistical techniques will be used where appropriate to quality assure the data input to CAMx. This method includes an analysis of the results from preprocessor programs.

6.1.1 Meteorology

Wind vectors, temperature, and the vertical exchange coefficient for each grid square will be plotted for selected hours and analyzed to determine if the data are appropriate, consistent, and correctly distributed.

6.1.2 Emissions Inventory

Daily emissions inventory summary graphics displaying grid cell emission densities for the various source types will be developed for each pollutant to determine if the emissions appear to be appropriate, consistent, and correctly distributed.

6.1.3 Air Quality

Air quality data for each monitoring location for selected hours will be plotted and analyzed to determine if the data are appropriate, consistent, and correctly distributed.

6.1.4 QA/QC Plan

The modeling staff conducts extensive Quality Assurance/Quality Control (QA/QC) activities when developing modeling inputs, running the model, and analyzing and interpreting the output. TCEQ has developed a number of innovative and highly effective QA/QC tools that are employed at key steps of the modeling process. Appendix C provides a detailed QA/QC plan developed by the modeling staff to be used during modeling for the Mid-Course Review and subsequently.

6.2 Diagnostic Testing of the Base Case Simulation

Diagnostic tests are designed to check the model's formulation and response to various inputs. Tests of this nature are routinely performed by the model developer and beta-testers when new models are first developed and released. CAMx is now a mature, well-tested model, not requiring this type of testing by end users in most cases. Therefore, the TCEQ performs only a limited number of diagnostic tests during the modeling process.

One test that we routinely perform is to compare the model output from different releases of CAMx with previous versions, using the same input data. After conducting head-to-head comparisons of CAMx Versions 3 and 4, we found some relatively minor - but somewhat unexpected - differences in the modeled output. Subsequent investigations by Environ, the developers of CAMx, showed a problem with the dry deposition algorithms that had been introduced in an earlier release. In response, Environ released the most current version of CAMx, version 4.03, which we are currently using in our modeling analyses.

6.3 Sensitivity Testing with the Base Case Simulations

Sensitivity tests are designed to check responses of the base case simulation to the plausible variability in the various model inputs. That is, given a possible change to some input parameter (e.g., doubling mobile emissions), the change in base case ozone production is determined. The results of these tests indicate the sensitivity of the model to various inputs and provide a guide by which modeling inputs may be reasonably adjusted to achieve acceptable model performance, as well as point out which inputs must be scrutinized most closely.

The following basic tests have been or will be performed to determine sensitivity to various model input parameters.

- *Alternative meteorological characterization* - Instead of the more common (but physically unreasonable) tests, such as one-half wind speed runs, we have instead tested CAMx using a variety of meteorological characterizations. The most significant of these is the use of the GOES-based meteorology during the core episode period, but we also tested the model with modified versions of the vertical mixing scheme. These tests helped identify the best meteorological characterization for use in this modeling application.
- *Alternative boundary conditions* - Because the area of interest is far from the lateral boundaries, it was expected that results will exhibit a great deal of sensitivity to the values. However, Environ ran tests of alternative boundary conditions as part of its modeling for the DFW area, and discovered a surprising level of sensitivity to the specification of boundary

conditions. We then conducted similar tests for the COMA, and discovered some sensitivity (though less than seen in the DFW modeling) to the boundary concentrations. Subsequently, we adopted the same boundary conditions as are currently being used in the DFW modeling.

- *Alternative emissions inventory assumptions* - Since the completion of modeling for the Phase 1 MCR, we have tested a number of different VOC adjustments using the Phase I MCR meteorology. As new analyses are completed, we intend to test the more promising VOC adjustment schemes in the hybrid base case. At the urging of Region VI, we plan to include a sensitivity run with adjustments to less-highly-reactive VOCs in the COMA.
- *Alternative vertical mixing* - EPA Guidance has historically recommended sensitivity testing to determine the model's response to perturbations in mixing height. Newer models, including CAMx, use instead a vertical mixing coefficient commonly known as K_v. In our application, the PBL derived from MM5 is used to calculate the K_v's used by CAMx. It is possible to test alternative PBL characterizations by modifying the PBL prior to the derivation of the K_v's. We already tested the profiler-adjusted PBL depths (described earlier) against the raw PBL depths output from MM5, and concluded that the profiler-based adjustment helped CAMx performance for days where the GOES nudging was not performed (no adjustment was seen to be necessary on the GOES days).

The base case has been established at this writing, with no major changes expected in the near future. Unless circumstances cause a substantial revision of the base case, no additional base-case sensitivities are planned.

In addition to the tests described in this section, tracer simulations were conducted to determine the contributions of initial and boundary conditions to the area of interest. These tracer simulations were described above in the section on boundary and initial conditions.

6.4 External Peer Review

The TCEQ contracted for independent peer review to be conducted for Phase 1 modeling of the August 22-September 1, 2000 TexAQS episode. This review was completed and the final report is available on the TCEQ web site. TCEQ has implemented or otherwise acknowledged all of the most significant suggestions provided by the Peer Reviewers. We also meet regularly with the Technical Committee and other stakeholder groups, and welcome their comments and suggestions towards how we can improve the modeling process.

7 Model Performance Evaluation

The performance of the base case modeling will be evaluated to determine whether the model is adequately simulating the formation of ozone. The model must show reasonable performance for each base case episode before the meteorological data for the episode are used with the future year emissions inventory to assess future control strategies.

At this writing, the base case has been established and evaluated using the usual one-hour analyses. We plan to conduct some of the additional performance analyses recommended by EPA for eight-hour modeling as well.

7.1 Performance Measures

Since the first day or two (ramp-up days) of a modeling episode are initializing days during which no (or minimal) exceedances were recorded, performance on these days is not considered in the overall evaluation of performance across an episode. Measures of performance include both qualitative (graphical) and quantitative (statistical) evaluations.

The photochemical model predicts a volumetric one-hour average over the whole grid cell. Monitoring data provides a measure of air quality at a specific point in space. To provide an accurate comparison with model predictions, the monitoring data would have to be transformed into volumetric one-hour averages over the same grid cells used in the model. However, monitoring networks are not dense enough to provide this information even for the most intensive studies that have been performed. Thus, comparison between the model's volumetric predictions and the monitored point measurements are the only recourse. This comparison can provide insight into model prediction trends but does not provide precise measures of model performance.
Additional information on specific procedures is found in the UAM modeling guidelines (U. S. EPA, 1991).

Some additional measures for evaluation model performance can be found in EPA's draft eight-hour modeling guidance which is available at the SCRAM website (U. S. EPA, 1999). The TCEQ plans to incorporate some of the recommended eight-hour performance measures into its evaluations but will continue to focus primarily on one-hour performance analysis, especially in the HGB area. The localized small-scale meteorological and emissions features characteristic of the area require model evaluation be performed at the highest temporal resolution possible to evaluate whether or not the model is "getting the right answer for the right reasons".

7.1.1 Graphical Methods

Graphical displays comparing predicted to observed concentrations can provide information on model performance. The following techniques will be used for days subsequent to the ramp-up day(s):

> Time-Series Plots - For each monitoring station in the domain and for each hour in the episode, the monitored concentration will be compared with the modeled concentration (interpolated from the four grid cell centers nearest the monitor). This comparison will determine whether the model can predict the peak concentrations and if the timing of ozone generation in the

model agrees with that found with the monitoring. Because modeled concentrations are compared with data from monitoring sites, which are specific points in space, it should not be expected that agreement would be excellent. Time series plots including the nine-cell minimum and maximum modeled concentrations will be produced to account for some of the inherent incommensurability between measurements and modeled concentrations;

Surface-Level Isopleths - For selected hours, surface-level isopleths (lines of equal concentration) will be drawn. This approach shows how the model is predicting the extent, location, and magnitude of ozone formation. This information can be compared to monitoring results;

Scatter Plots - Scatter plots of predictions compared to observations depict the extent of bias in the ensemble of hourly data pairs. Systematic positioning of data points around the perfect correlation line indicates bias. The distribution of points over the area is an indication of error. This procedure also indicates outlier pairs;

Animations - Model output will be rendered into an animated sequence showing the formation and transport of ozone (and its precursors) throughout each episode. These animations will be compared to the conceptual models developed for the respective episodes to assure that the model replicates TCEQ’s understanding of the process. If the animation differs fundamentally from the conceptual model, then both the model formulation and the conceptual model will be reviewed and revised as appropriate; and

Aloft measurements - During the TexAQS 2000, numerous aircraft flights collected a rich set of above-ground ozone, ozone precursor, and reaction product measurements. Additionally, data was collected at the Williams Tower at an elevation of approximately 250 meters above ground level (AGL). Data from aircraft transects will be compared with model predictions along the flight path. Data collected at the Williams Tower will be compared with model predictions at the appropriate vertical layer using time-series plots as described above.

<u>7.1.2 Statistical Methods</u>

These methods can provide a quantitative measure of model performance. The results must be considered carefully, especially in cases where there are not a large number of monitors. EPA recommends the following statistics for use in evaluating performance of the model for one-hour attainment demonstrations (U. S. EPA, 1991):

- Unpaired Highest-Prediction (Peak Domain Maximum) Test: This measure compares the difference between the highest observed value and the highest predicted value found over all hours and over all monitoring stations. We will make this comparison for both one- and eight-hour peak ozone concentrations.

- Normalized Bias Test: This test measures the model's ability to replicate observed patterns. Since there are many time periods when relatively low levels of ozone are predicted and statistics from these periods are not very meaningful, this test will be limited to pairs where the observed concentration is greater than 0.060 parts per million (ppm). This threshold is notably above the naturally occurring ozone background value of 0.040 ppm.

- Gross Error Test: This test compares the differences between all pairs of predictions and observations that are greater than 0.060 ppm. This examination is a measure of model precision.

7.1.3 Eight-Hour Performance Metrics

In its Draft Guidance (U.S. EPA, 1999) the EPA also recommends several metrics for use in evaluating the performance of the model relative to eight-hour ozone prediction. Unlike the one-hour statistics, which are based on multiple observation-prediction pairs per day (whenever the observations are > 60 ppb), all the eight-hour metrics are based on comparing the daily peak observed eight-hour concentrations with peak eight-hour modeled concentrations near (i.e. within a few grid cells of) the monitors. Because these metrics are focused on only the daily peak-to-peak comparison, they are not as sensitive to some types of prediction errors as the traditional one-hour performance measures. Because of this, the TCEQ will continue to primarily assess model performance based on one-hour statistics, isopleth plots, and time series, although we will calculate two of the recommended eight-hour metrics as well. These are:

- Bias - the tendency of the model to over- or underpredict the monitored eight-hour peaks

- Correlation - The overall correspondence between measured and modeled peaks.

The Draft Guidance also recommends calculating a “Temporal Correlation” measure as well as producing quantile-quantile plots and fractional bias calculations for various subsets of the measured-monitored data pairs. We will investigate the usefulness of these metrics, but may not include them in the final SIP revision if they are not found to be insightful or are confusing.

7.2 Assessing Model Performance Results

Model performance evaluation tests are performed for each base case for days subsequent to the ramp-up day(s). The goals for the results of the one-hour statistical tests are the following:

Unpaired peak prediction: +/- 15 - 20 percent;
Normalized bias: 5 - 15 percent; and
Gross error: +/- 30 -35 percent.

For the eight-hour metrics, the EPA recommends a tolerance of +/- 20% for the bias and for the relative bias tests. No recommended ranges for the other recommended eight-hour metrics are provided.

If the statistical measures for a base case do not fall into recommended ranges or if the graphical analysis indicates poor performance, the input data for the base case are carefully analyzed along with the results of the sensitivity tests. When appropriate, certain sensitive inputs developed with uncertain data may be modified to yield better model performance. Any modifications to input data are always coordinated with the Photochemical Modeling Technical Committee and the EPA Region 6 office. This process must be approached very carefully because good model performance must be obtained for the correct reasons and must not be considered an end goal in itself.

7.3 Ozone Precursors and Photochemical Products

One of the most important uses of the TexAQS 2000 data is to provide model comparisons against species not routinely monitored, such as formaldehyde, nitric acid, and methylperoxyacryloyl nitrate (MPAN), and other similar alkyl nitrates. Comparing the model predictions against these data can provide great insight into the modeled emissions and the model's chemical processes[1]. TCEQ has made comparisons between most modeled photochemical reaction products and their measured counterparts, as well as having compared ozone precursors such as NO, NO_2, and various hydrocarbon species with measurement data. The preferred analysis technique for these comparisons is through examination of time series plots.

While performance analysis of non-ozone species is very valuable and provides great insight into the model's workings, it is not appropriate to base model performance evaluations directly on these species. First and foremost, CAMx (as well as similar models) is optimized to predict ozone, not NO, NO_2, PAN, or any other CB-IV species besides ozone. Second, many emissions of primary species are on a scale much smaller than the model's finest spatial resolution. Ozone, on the other hand, is a secondary pollutant and its concentration normally is expected to remain relatively constant across areas of a few to several kilometers in width (one reason why short-term ozone peaks are difficult for the model to replicate). Finally, no statistical performance evaluation criteria are available from EPA for non-ozone species, so only graphical performance analysis techniques can be applied.

Nonetheless, TCEQ carefully considers how well the model replicates the spatial and temporal distributions of all CB-IV species for which comparable measurements are available. Major discrepancies involving these species are investigated to seek causes in the model formulation.

7.4 Subregional Analysis

Most model comparisons, with the exception of the EPA Performance Criteria, are actually performed on a subregional basis. Time series plots compare model performance against individual monitors, and ozone isopleth plots are compared with monitored values at discrete locations. Some modeling applications have divided the domain into smaller regions and applied the EPA Performance Criteria

---

[1]MPAN is not presently included in the CB-IV mechanism, but can be used to evaluate modeled emissions of isoprene.

within those regions. However, this approach is not always useful and could even be abused if the subregions are not defined objectively. TCEQ has avoided arbitrarily subdividing the modeling domain in the past (except for obvious divisions, such as HGB vs. Beaumont-Port Arthur) since there has been no particular rationale for carving up the domain.

Recently, however, a great deal of attention has fallen on the possibility of distinguishing routine urban ozone formation from that seen downwind from the heavily industrialized areas. TCEQ is investigating ways to divide the HGB in such a way as to be able to assess model performance in the "routine urban" ozone plume separately from the industrial plumes which occur in the area, especially the plume emanating from the Ship Channel. This analysis will be complicated by changing wind directions and the fact that sometimes the plumes combine.

In addition to the analyses targeted at resolving the urban and industrial plumes, TCEQ also plans to conduct tests to search for systematic spatial biases in the model predictions. One possible test plots the residuals (modeled - measured concentration) in the 4 km sub domain as a function of latitude and another will plot the residuals as a function of longitude. These tests should help to identify any systematic bias resulting from location.

7.5 Conceptual Model Evaluation

The model is evaluated against the day-specific conceptual model (Appendix B). In general, the model should accurately represent the physical phenomena known to occur and contribute to ozone formation, such as the onset of the sea breeze and the daily clockwise veering wind pattern now known to be associated with ozone exceedances in the region. Once the most significant physical phenomena have been identified for each episode, the modeling staff will devise a series of tests, both quantitative and qualitative, designed to measure how well the modeling system replicates these phenomena. Tests developed for modeling studies in the San Joaquin Valley of California provide examples of the types of tests that will be developed.

7.6 Process Analysis

Recently Process Analysis (PA) was added to the CAMx in version 3.10, which was released in March, 2002. This extremely valuable tool lets modelers study the intimate details of ozone formation, showing the various physical and chemical processes that determine the modeled ozone concentrations at specified locations and times.

If model base case performance does not prove to be acceptable, TCEQ makes extensive use of process analysis to try to determine the causes. Even if model performance is good overall, TCEQ uses process analysis to help ensure that the model is "getting the right answer for the right reasons". The modeling staff received training in late March, 2002 and has applied process analysis to early modeling runs. TCEQ is continuing to apply process analysis to selected modeling runs as they are conducted.

7.7 Policy Implications of Model Performance

While there are currently no procedures in place for considering a model's performance in policy-making, it is an important issue which should be acknowledged. The model provides a deterministic "pass-fail" test for demonstrating attainment, while in actuality attainment is determined based on ambient concentrations which vary according to some probabilistic distribution. This random component is further complicated by the fact that the model never provides an exact replication of reality. Models inherently possess some level of uncertainty, typically measured as bias and gross error, which should factor into the decision-making process. As model performance improves, greater reliance can be placed on modeling, but no formal process for quantifying this relationship is currently available. Future developments may provide such a process, but at present it is very important to consider overall model performance in factoring model results into policy decisions.

8 Attainment Year Strategies

The modeling staff will continue to work with Policy to conduct runs proposed by Policy staff in a timely manner. Changes to the future case will be based on a reevaluation of local, state, and federal rules. After adopted rules are incorporated, various strategies may be modeled.

9 Procedures to Archive and Document Study Results

EPA recommends that certain types of documentation be provided along with a photochemical modeling attainment demonstration. TCEQ is committed to supplying the material needed to ensure that the technical support for any SIP revision is understood by all involved parties. To that purpose, TCEQ will document the following items in conjunction with the attainment demonstration:

> Modeling Protocol - Establishes the scope of the analysis and encourages stakeholder participation in both the study development and the study itself;
>
> Emissions Inventory Final Report - Summarizes the development of the model-ready emissions estimates database. This report will contain tabular and graphical summaries of the data for both base and future years;
>
> Air Quality/Meteorological Input Final Report - Summarizes the development of the meteorological and other needed model input fields. This report will contain tabular and graphical summaries of the relevant data;
>
> Model Performance Evaluation Report - As discussed in Section 7, an assessment of the suitability of the model to support emissions control policy will be assessed. The findings of that analysis will be discussed comprehensively in the model performance evaluation report. Also, as discussed in Section 6, several diagnostic analyses are planned to determine whether the photochemical modeling results are physically sound. The results from these analyses will be included as part of the performance evaluation report;

Description of the Attainment Demonstration Strategy - The documentation (likely as part of a final report) will outline the specific control measures which embody the attainment demonstration plan. A description of the modeling, which suggests attainment will be achieved in a future year, will be provided. If any "weight-of-evidence" arguments are used to supplement the findings of the air quality modeling, a description of the techniques used and a summary of the findings will also be documented;

Graphical depictions of the modeling results - Ozone isopleth plots, difference isopleth plots, and ozone animation sequences, will be produced to aid in sharing model results with EPA, TCEQ management, and stakeholders; and

External Review - TCEQ will document the review procedures (internal and external) employed in the project. This approach will include instructions for interested external parties to access the study database, including software utilized as part of the technical analyses.

Note that the above list is not all-inclusive and that additional documentation will likely be developed in the course of fully documenting the modeling activities. Some items may be documented as part of the actual SIP, while others will be provided as Appendices, Attachments, or Supplementary Reports. All relevant documentation will be available electronically, either through the TCEQ web site www.tceq.state.tx.us or by contacting TCEQ.

As per the request of EPA Region 6, the documentation package will contain several new or expanded discussions, including the following specific topics:

- Point source speciation
- Methods and rationale for adjusting both HRVOCs and less-reactive VOCs
- HRVOC cap determination for control strategy purposes
- Growth and controls for sources outside the HGB nonattainment area.

TCEQ will also archive all documentation and modeling input/output files generated as part of the COMA modeling. Dr. Jim Smith of TCEQ will be responsible for these products and may be reached by telephone at (512) 239-1941 or via e-mail, jismith@tceq.state.tx.us for information regarding data access or project documentation.

10 Bibliography

Guenther et al., 1999. Isoprene emission estimates and uncertainties for the Central African EXPRESSO study domain. *J. Geophys. Res.* 104(D23): 30,625-30,640.

Guenther et al., 2002. *Biogenic VOC emission estimates for the TexAQS 2000 emission inventory: Estimating emissions during periods of drought and prolonged high temperatures and developing GloBEIS3.* Final report. Prepared for Mark Estes, TNRCC, April 2, 2002.

Kinnee et al., 1997. United States land use inventory for estimating biogenic ozone precursor emissions. *Ecological Applications* 7(1): 46-58.

Mendoza-Dominguez et al., 2000. Modeling and direct sensitivity analysis of biogenic emissions impacts on regional ozone formation in the Mexico-United States border area. *J. Air & Waste Manage. Assoc.* 50: 21-31.

O'Brien, 1970. Journal of the Atmospheric Sciences, v. 27, 1970

Pinker, Rachel, 2002. *High resolution solar radiation data for biogenic emissions modeling for 2000 ozone episodes in the Houston area.* Final report. Prepared for TNRCC, August 30, 2002.

Tanaka, P. L., S. Oldfield, J. D. Neece, C. B. Mullins, D. T. Allen, Anthropogenic sources of chlorine and ozone formation in urban atmospheres. *Environ. Sci. Technol.* **34**, 4470-4473 (2000). http://pubs.acs.org/subscribe/journals/esthag/jtoc.cgi?esthag/34/21

Tanaka, Paul L., et. al., Direct Evidence for Chlorine-Enhanced Urban Ozone Formation in Houston, TX, submitted to Atmospheric Environment, April, 2002.

Tanaka, Paul L., et. al., Development of a chlorine mechanism for use in the CAMx regional photochemical model, submitted to J. Geophys. Res., April 2002

Tanaka, Paul L., Charles B. Mullins, and David T. Allen, An Environmental Chamber Investigation of Chlorine-Enhanced Ozone Formation in Houston, TX, submitted to *J. Geophys. Res.*, April, 2002.

Tonnessen, Gail S., Process analysis of Houston SIP Modeling, available on a CE-CERT web page, http://pah.cert.ucr.edu/hpa

U.S. EPA, 1991. Guideline for Regulatory Application of the Urban Airshed Model

U.S. EPA, 1999. Draft Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-Hour NAAQS

Vizuete, et al., 2002. Effects of temperature and land use on predictions of biogenic emissions in Eastern Texas, USA. *Atmospheric Environment* 36(20): 3321-3337.

Wiedinmyer et al., 2000. Biogenic hydrocarbon emission estimates for North Central Texas. *Atmos. Environ*. 34: 3419-3435.

Wiedinmyer et al., 2001. A land use database and examples of biogenic isoprene emission estimates for the state of Texas, USA. *Atmos. Environ*. 35: 6465-6477.

Yarwood et al., 1999. *Development of Globeis—A state of the science biogenic emissions modeling system.* Prepared for Mark Estes, TCEQ, December 23, 1999. Pp. 103.

Yarwood et al., 2001. *Biogenic emission inventories for regional modeling of 1999 ozone episodes in Texas.* Prepared for Mark Estes, TCEQ, March 30, 2001. Pp. 52.

Additional References

Summary of Chlorine Chemistry Studies.
http://www.tnrcc.state.tx.us/air/aqp/airquality_contracts.html#other02

Spatial and Temporal Impacts of Chlorine Chemistry on Ozone Formation in Southeastern Texas
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/Impacts_Chlorine_on_Ozone_Formation.pdf

Impact of Chlorine on Ozone Modeling for the Houston Area.
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/ImpactOfChlorineOnOzoneModeling.pdf

Confirming the Presence and Extent of Oxidation by CI in the Houston Texas Urban Area Using Specific Isoprene Oxidation Products as Tracers.
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/ConfirmingPresenceandExtentOfOxidationByCI.pdf

Impact of Molecular Chlorine Emissions on Ozone Formation in Southeast Houston.
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/ImpactOfMolecularChlorineEmissionsOnOzoneFormation.pdf

Incorporation of Chlorine Reactions into the Carbon Bond IV Mechanism.
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/IncorporationOfChlorineReactionsIntoCB4-1.pdf

Incorporation of Chlorine Reactions into the Carbon Bond IV Mechanism: Mechanism Updates and Preliminary Performance Evaluation.
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/IncorporationOfChlorineReactionsIntoCB4-2.pdf

Emission Inventory for Atomic Chlorine Precursors in Southeast Texas.
ftp://ftp.TCEQ.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/Others/EmissionInventoryForAtomicChlorinePrecursors.pdf

Appendix A

Conceptual Model for Ozone Formation in the Houston/Galveston Area

available at:
http://www.tnrcc.state.tx.us/air/aqp/airquality_photomod.html#tsd2

Appendix B

Meteorological and Ozone Characteristics in the Houston Area from August 23 through September 1, 2000

available at:
http://www.tnrcc.state.tx.us/air/aqp/airquality_contracts.html#aq04

Appendix C - Photochemical Modeling QA/QC Plan

Photochemical Modeling QA/QC Plan

In order to ensure that the inputs to the CAMx model are of the highest possible quality, the TCEQ Air Modeling Team performs a series of rigorous quality assurance procedures on the model input files. All data produced by external contractors are examined by a member of the air modeling staff.
Biogenic Emissions Estimates

Biogenic emissions estimates are produced using GloBEIS, which requires temperature data and photo synthetically active radiation data as inputs. Measured temperature data taken from the CAMS network, offshore buoys, an agricultural temperature network, and the National Weather Service are interpolated into hourly temperature fields. Contour plots are then made of the hourly fields to check for reasonableness. Desirable features of the data include:

Small diurnal temperature variations over the Gulf of Mexico;

Larger diurnal temperature variations over land;

Differences between temperatures in rural and urban areas. Urban areas should have higher overall temperatures; and

The effect of rain on the temperature. Any rain storms should decrease overall temperatures.

Photosynthetically active radiation data, the second type of input to the GloBEIS model, are extracted from satellite data. To qualitatively check the extracted data, they are compared to total solar radiation values measured at ground-based sensors. The time variation of the extracted data is compared to the time variation of the measured data. The two time variations should exhibit the same pattern.

In addition to the qualitative check above, a quantitative check is performed. Total solar radiation measured at a ground-based sensor is plotted against the derived photo synthetically active radiation for that same area. The slope of the line produced is usually close to 0.5, and the correlation coefficient is usually high, which indicates that the derived data is successfully picking up cloud fields.

Once the biogenic emissions estimates for every hour of the ozone episode have been produced by GloBEIS, qualitative checks are performed on them by creating tile plots and time series. Anomalies in the data are sought, including:

The timing of isoprene production, i.e., does isoprene production begin at sunrise and end at sunset?

Unexpected geographic allocation of emissions, e.g., are there emissions located over the bay or high isoprene production in downtown Houston?

To document quality assurance activities, a log file is created for each GloBEIS run. This log file contains the name and location of the input and output files, the date of the modeling run, the operator name, and a brief description of the run. Further file traceability is provided by the tileplots produced, which are stamped with the name and date of the original file. This policy provides a second QA person the opportunity to trace file names from the tileplot back through the log file to the original input files.

Quantitative comparisons of the modeled data to measured data are documented by producing time series, scatterplots with regression statistics, and performance statistics. All graphs and statistics are stamped with the file name, date, location, and the date the plot was made.

Onroad Mobile Emissions Estimates

Estimates of onroad mobile emissions are produced by the Texas Transportation Institute (TTI). The data files received from TTI are processed in SAS to convert the times from daylight savings to standard time. Then the data are processed by three modules of EPS 2x: LBase, ChmSplt, and GRDEM. The resulting data files are then quality assured to ensure that the right amount of emissions are located in the right place (spatial allocation) at the right time (temporal allocation).
TTI provides summary files which have emissions totals for all pollutants for the entire nonattainment area as well as for each individual county. To ensure that no emissions are lost in the EPS 2x processing steps, the emissions totals for each pollutant in the message files of LBase, ChmSplit, and GRDEM are compared to the totals in the TTI summary files. First, the totals for the nonattainment area as a whole are compared to ensure that no gross emissions have been omitted. Then the totals for each county are compared.

To check the spatial and temporal allocation of the emissions, the data are fed into a tile plot program. The resulting plots are then checked to ensure that the emissions fall along the major roadways. In addition, the diurnal profile is checked to see if it follows the “Batman curve” (a curve with two peaks around 8 am and 5 p.m.) on Mondays through Thursdays.

Area Source Emissions Estimates

The processing of area source emissions is automated via UNIX scripts. First, the area source surrogates are tested for validity. Ten tons of test emissions are allocated to each county within the nonattainment area and viewed with tile plots to check that they are reasonably distributed throughout the counties. Second, the actual emissions are divided into various categories (e.g., logging, lawn and garden equipment, offshore shipping, etc.) by a PERL script. Third, a UNIX shell script is used to process the emissions through EPS 2x. Each category of emissions is processed separately. Once the emissions are gridded, each category is viewed using tileplots to check for reasonableness.

Quality assurance of the emissions is also automated. One program deletes all the intermediate EPS 2x files and confirms that EPS 2x generated no significant errors. Any error messages generated by EPS 2x are written to the program output. A second program tracks the emissions amounts through the EPS 2x processing steps. It summarizes the input and output emissions for every processing step, showing where any emissions were lost and why.

Point Source Emissions

Point source data are extracted from the TCEQ's Point Source Database (PSDB). Many QA/QC checks are performed on the PSDB data that are extracted. Initially, the annual emissions totals by pollutant, stack, SIC, and SCC are compared to the respective average of the previous five years. The annual emissions should generally be within three standard deviations of the five-year average. Any outliers that are not easily explained are reported to the Industrial Emissions Assessment Section (IEAS) team leaders for correction or explanation. Next, the stacks are checked for outlier stack parameters. The parameters of each stack are compared to the average of the previous five years to ensure that they are within three standard deviations of the five-year average for that type of facility (SIC and SCC). Again, any outliers that are not easily explained are reported to the IEAS team leaders for correction or explanation. Default stack parameters by source type are used to substitute for outlying stack parameters. Next, the data records are checked for incorrect/outlier stack coordinates (geographical locations). If a stack's coordinates are located outside of the county where the plant is located, they are changed to the coordinates of the center of the county. Default start times for selected pieces of equipment are substituted for equipment with faulty start times. There is also a check performed for missing SAROAD/AIRS pollutant codes, among many other small QA/QC checks performed directly on the extracted PSDB data.

Next, the overall performance of the modeling extract is checked against "Top 100" Paradox queries of the PSDB performed by the IEAS. The top 100 NOx-emitting stacks in the state, from the modeling extract, are compared to a similar list obtained from IEAS queries. This process is performed to ensure that no large emitters of NOx have been omitted during the extraction process or subsequent PSDB QA/QC checks. A similar check is performed for the top 100 VOC-emitting stacks, and any discrepancies between the two lists are reconciled prior to any further point source processing. A comparison of the "Top 100" lists with previous "Top 100" lists is performed, to ensure that large emitters in the state have not been accidentally dropped off the lists and to ensure that any new large emitters in the state are correctly on the lists.

The last step of the PSDB extract and QA/QC processing is the creation of AIRS Facility Subsystem (AFS) records, as required for EPS2x preprocessing. The AFS records for the state are then split into electrical generating utilities (EGU) and non-EGU (NEGU) AFS files. The primary reason for splitting the AFS file into EGU and NEGU AFS files is to ease facilitation of incorporating hourly CEM data from EPA's Acid Rain Program Database (ARPDB) into the EGU AFS file. The PSDB-to-ARPDB cross-reference file is updated each time a new modeling episode is chosen to ensure the best link between the PSDB stack identifiers and the ARPDB identifiers. The cross-reference is double-checked to ensure that every boiler in the ARPDB data matches a FIN/EPN (emission point/stack) in the PSDB data. This cross-reference must be used in the merge of the ozone-season daily PSDB data with the hourly ARPDB data. After the merge, a manual sampling of records is performed to make sure that there is a record for every pollutant for every hour of every day of the ozone episode, for each EGU in the ARPDB.

Once hourly records from the ARPDB records have replaced the corresponding PSDB records, at least two sources are chosen at random for quality assurance tracking. Two days of the episode are chosen at random, and the emissions of NOx (and SO_2, if SO_2 is being modeled) for every hour are compared to their respective values in the original ARPDB data to ensure that no emissions were lost or gained during processing. The emissions of two potential problem (mis-allocation of emissions from ARPDB boilers to PSDB stacks) stacks, such as common stacks and multiple stacks, are also specifically checked. Hourly VOC and CO emissions for each ARPDB stack are allocated according to the hourly NOx profile of each day for each ARPDB stack. The calculated VOC and CO emissions are checked to ensure that the program has allocated them completely and accurately, and that the PSDB NOx/VOC and NOx/CO ratios are maintained for every hour of the day.

After the EGU data have been thoroughly checked, they are processed through EPS2x. “Tileplots” are created for the gridded low-level emissions and elevated emissions to facilitate additional QA/QC of spatial and temporal allocation. Consistent with electrical usage patterns, EGUs usually produce a diurnal profile with a broad peak near mid-afternoon. More details about tileplots as a QA/QC tool are given below.

If hourly Special Inventory (SI) data is available, incorporation of SI data follows a very similar QA/QC routine as the ARPDB data. SI data may be applied to both the EGU AFS file and the NEGU AFS file. AFS files are created for all states within the modeling domain, and ozone season or annual emissions records are given very nearly the same thoroughness of QA/QC as the state of Texas and the nonattainment area. When hourly data for other states is incorporated into those ozone season or annual records, the same level of QA/QC is applied as for Texas processing, to ensure that emissions are not lost or gained, and that the files are as accurate as the available raw data. A SAS program, sum_afs, is executed to compare the overall ozone season daily and hourly emissions for a day’s worth of AFS records with that day’s output from the EPS2x PREPNT module.

As with area sources, the processing of point source emissions through EPS2x is fairly automated via UNIX scripts. The point source emissions for the entire domain are typically divided into manageable pieces. Within each piece of the modeling domain, the data are completely related in that they have either come from the same data source, or the data will be expected to have similar growth and controls applied to them for the future case. For example, in recent modeling studies, the pieces included: Texas EGUs, Texas NEGUs, Louisiana EGUs, Louisiana NEGUs, Offshore, Mexico, Regional EGUs, and Regional NEGUs. The level of QA/QC drops off only slightly with distance from the nonattainment area.

QA/QC of the EPS2x processing steps is also fairly automated. Scripts are written to capture any error messages generated by EPS2x at each processing step.
A tracking summary program combines all of the message file input and output emissions values from each EPS2x processor program. This tracking program generates a single file for easy QA/QC of emissions gain and loss from a given portion of the inventory. As stated above, tileplots are generated for each low-level and elevated component of each portion of the inventory.
These steps are repeated for each piece of the modeling domain as it is processed through EPS2x. Low-level pieces of the point source inventory are merged together via EPS2x module MRGUAM. The input to MRGUAM is a list of the pieces to be included in each model run, so this list is QA-ed thoroughly. A summary output from MRGUAM of all of the emissions from each piece can be

compared to previous or similar merges to ensure that emissions are as expected. The low-level total emissions for each day is also provided as output from this step. As a final QA step for low-level sources, a tileplot of the total merged points is generated. The typical diurnal profile for low-level point sources is fairly flat across the day, with typical variations across hours being less that ten percent.

A similar process is performed for elevated point sources prior to input to the photochemical model. Since elevated sources generally only come from point sources, any change in elevated emissions is due to point sources. Tileplots are generated for every merged elevated file, the spatial distribution is compared to previous merges, the overall emissions total on the tileplot is verified, and the diurnal profile for each day is checked. The typical diurnal profile of elevated point sources peaks in mid afternoon because EGU emissions are the dominant source type in the elevated files.

MM5 QA/QC Procedures

Quality checking meteorological fields for MM5 requires several pre-processing steps. The output from each pre-processor provides a check for some of the variables and parameters which are part of MM5 input fields. Running MM5 itself requires verification through the job deck, mm5.deck, that switches and options have been correctly selected. The quality assurance of MM5 output is a central part of the technical work plan designed by TCEQ staff and Dr. Nielsen-Gammon of Texas A&M University and is intended to provide a rationale for the sensitivity analyses which followed the initial modeling.

A description of each of the pre-processor configurations was provided in the report *Initial Modeling of the August 2000 Houston-Galveston Ozone Episode, December 19, 2001*. Of particular importance was the evaluation and graphical inspection of surface characteristics and sea temperatures after running TERRAIN, PREGRID, and REGRIDDER. The "basic-state" of the Houston atmosphere, as characterized by the sea level temperature and pressure, lapse rate, and stratospheric isothermal temperature, was selected by Dr. Nielsen-Gammon to reflect the Houston summertime environment in the processor INTERPF.

Due to the importance of the land use and water characteristics on forcing the MM5 surface boundary, special attention was directed towards a realistic specification of key surface parameters. In particular, the adjustment of soil moisture availability was made by referencing climatological data, and in part by analyzing the Bowen ratio. These are referenced in more detail in the above referenced report.

The availability of global analysis fields and observational data is potentially very useful for model nudging. However, the use of data which has not already been checked for quality could destabilize MM5 or produce unrealistic responses which do not accurately reflect the true atmosphere. The global analysis fields which were used for model initialization and boundary conditions already have undergone review by NCEP modeling staff. Use of these fields for analysis nudging allowed for further inspection. Surface observations were collected and inspected by Dr. Nielsen-Gammon under a contract completed prior to MM5 modeling. Review and quality assurance of GPS sonde, rawinsonde, and profiler data was performed by NOAA Environmental Technology Laboratory and Aeronomy Laboratory staff. The completed document will be available by August 31, 2002.

Quality Assurance

The use of generic input and output file names is part of somewhat complicated MM5 and pre-processor job decks, so separate output directories were created for each run to avoid over writing and to help manage post process analysis. Job decks were preserved in these directories, and a log of model runs was maintained at www.met.tamu.edu/results. A separate file was maintained on the TCEQ SGI Origin 3400 ("typhoon") labeled "state of the art model" to document the "best" current configuration of the MM5 model output to be used for SIP modeling.

Review of the job decks for MM5 and each of the pre-processors was accomplished by TCEQ staff. In addition to reports provided to TCEQ and associated WWW links, Dr. Nielsen-Gammon and TCEQ staff presented summary reports at Interim Science Coordinating Committee meetings and at Technical Review Committee meetings to members of the broader scientific community who participated in the TexAQS 2000 data intensive as well as stakeholders who have had long standing familiarity with meteorological modeling issues for the Houston-Galveston area.

Input File Tracking

TCEQ uses the air quality modeling log database to log and track the ozone photochemical modeling runs. The database contains the information of the air quality model used for a run, modeling input files and descriptions, job control or script file, and QA information. The database is implemented with the MySQL database system. The web browser based user interface was created to allow the modeling staff to make log record entry and query. The log database record fields are listed in the following table:

Air Quality Modeling Log Database

Item	Type	Description
id	int	unique ID number of a modeling run (used internally by the database)
log_date	date	date of the modeling run and logging
log_name	char	name of staff who makes the run and log entry
qa_date	date	date of QA is conducted
qa_name	char	name of staff who QAs the run
project	char	project name, e.g. HGMCR
episode	char	episode dates, e.g. 20000822-20000901
model_type	char	name of AQ model, e.g. CAMx
model_version	char	version of AQ model, e.g. 3.10
case	char	modeling case: base or future and version #, e.g. base1 (part of output file name)
scenario	char	sensitivity or control scenario, e.g. 070wind for 70% wind (part of output file name)
el_ei_type	char	case and scenario of elevated EI (part of input file name), e.g. base1.regular
lo_ei_type	char	case and scenario of low-level EI (part of input file name), e.g. base1.regular
met	text	description of meteorological inputs
ei_bio	text	description of biogenic EI inputs
ei_area	text	description of area and non-road mobile EI inputs
ei_point	text	description of popint source EI inputs
ei_mobile	text	description of on-road mobile EI inputs
other	text	description of other features
job	text	modeling job control files/script

This database is implemented with MySQL with the web-browser-based user interface for log entry and query.